A Powerful Tool for Fitting and Forecasting Deterministic and Stochastic Processes : the Kohonen Classification

Eric de Bodt[1], Philippe Grégoire[2], Marie Cottrell[3]

[1]Université de Lille 2 / Université catholique de Louvain,
[2]University of Namur – CeReFim, [3]Université de Paris 1 – SAMOS

Abstract. In this paper, we propose a general approach for fitting and forecasting the behavior of time-dependent processes. The only hypothesis on which it is based is the stationarity of the process dynamics. The approach is clearly non-parametric, uses no kind of a priori hypothesis on the form of the process and reveals itself powerful on either deterministic processes (such linear, logarithmic or sinusoidal ones) or stochastic ones (being able to reproduce even a white noise). The fields of applications of the proposed methods are time-series prevision but also risk analysis, allowing to determine the limits between which a stochastic process will behave on a specific time-horizon.

1. Introduction

The method that we propose in this paper is oriented towards time-series analysis. Its goal is to capture the dynamics of a process observed through a set of data to simulate its future behavior. The only hypothesis on which it is based is the stationarity of process dynamics during the period of analysis. The main features of our approach are its *non-parametric nature* (there is no kind of hypothesis about the specific form of the analyzed process), its ability to incorporate *multi-dimensional times series* (presented as vectors) and to cope either with *deterministic processes* or *noisy ones* (even with white noise, as we will show below). The main fields of application of this approach is, clearly, *time-series prevision* but also *risk analysis*, putting into light the limits into which the analysed process should be contained during a specific time horizon. The proposed approach certainly has relations with other non-parametric regression approaches (see for example [9]) but this remains to be studied.

In the second section of the paper, we present in few words the Kohonen algorithm, used here as a vector quantization algorithm for the discretization of the process dynamics. In the third section, we describe the proposed method. In the fourth one, we show applications to deterministic processes (linear, logarithmic and sinusoidal), to a white noise and to a stochastic process (well-know in the field of interest rate structure modelisation).

2. The Kohonen Algorithm

The Kohonen algorithm [7, 1, 2] is a well-known unsupervised learning algorithm which produces a map composed by a fixed number of units. A physical

neighborhood relation between the units is defined and for each unit i, $V_{r(i)}$ represents the neighborhood with radius r centered at i. Each unit is characterized by a parameter vector W_i of the same dimension as the input space. After learning, each unit represents a group of individuals with similar features (this group is named Voronoi region of the unit). The correspondence between the features and the units (more or less) respects the input space topology : similar features correspond to the same unit or to neighboring units. The final map is said to be a self-organized map which preserves the topology of the input space.

While the asymptotic properties of this algorithm remain partly unknown, some of its theoretical properties have been demonstrated during the last 10 years. One of them is of particular interest as regards this paper and concerns its density approximation property. In [8], the Kohonen algorithm terminating with a 0 neighbour at the end of learning (that is the classical competitive quantization) is studied. The author shows that the units after learning are a good discrete skeleton for reconstructing the initial density f, provided that each unit is weighted by the probability (estimated by the frequency) of its Voronoi region. In other terms, if y_1, $y_2,\ldots,y_n$ are the code vectors after learning, and C_1, $C_2,\ldots,C_n$ the corresponding Voronoi regions and P the initial probability distribution associate to density f, the following convergence (in law) is guaranteed:

$$\sum_{i=1}^{n} P(C_i)\,d_{y_i} \xrightarrow{\text{law}} P \qquad (3)$$

when n goes on to infinity, and d_{y_i} is a Dirac function on y_i. This is equivalent to say that the empirical measurement defined by units y_1, $y_2,\ldots,y_n$, weighted by the probabilities of the associated Voronoi regions, converges (in law) on the initial probability P. This remarkable propety is true because with zero neighbor, the Kohonen algorithm is nothing else than an usual Competitive Quantization. But actually as the classes that we obtain are topologically ordered, the can be represented in a convenient way and they can be easily grouped if necessary. Moreover, the convergence is accelerated by the non zero neighbor phases. For more information on those theoretical aspects, see [3, 5].

3. Fitting and Forecasting Processes

To fit a particular process, we start from a data matrix composed by observations of the process at different points in time (each row correspond to a point in time). The process can be characterized by one or several variables. The initial data matrix, denoted D, is therefore of order [r x c], where r is the number of observations in time and c the number of observed variables.

The first step of our approach is, as for a lot of classical time series analysis methods, to choose a lagging order. The initial data matrix, D, is then modified to incorporate, for each row, the vector of the observed variables as well as the past realizations of those ones. The new data matrix, LD, is of order $[(r-\lambda) \times (c \times \lambda)]$,

where λ is the lagging order. Rows of LD are denoted $x_t = \{x_{t,1}, \ldots x_{t,n}\}$, where t is the time index and $n \in (c \times \lambda)$.

The LD matrix is then decomposed into a number of homogeneous clusters, using the Kohonen algorithm (a one dimensional Kohonen map is used). To each unit of the Kohonen map is associated, after learning, a number of individuals for which this unit is the winning one. The clusters are formed. For each unit, the mean profile ($\overline{x}_i$) of the attached individuals is calculated. As λ, the choice of the number of clusters will depend on the features of the analyzed process. The observation of the homogeneity of the clusters (for example, the Fisher statistics or one of its multidimensional extensions) obtained after learning is a good indicator.

For each row x_t of the LD matrix, the associated deformation is then computed. It is denoted y_t and is obtained by the following calculation : $y_t = x_{t+t} - x_t$, where τ is a time delay. On this basis, for each cluster of the LD matrix, a P_i matrix, composed by the y_t corresponding to the x_t of the cluster i, is formed.

For each P_i matrix, as for the LD matrix, a decomposition into a number of homogeneous clusters is realized, once again using the Kohonen algorithm (a one dimensional Kohonen map is also used). The mean profiles of the formed clusters are then determined and denoted by $\overline{y}_j$.

The last step to characterize the analyzed process is to establish the empirical frequencies of $\overline{y}_j$ conditionally to $\overline{x}_i$. They are denoted by $P(\overline{y}_j | \overline{x}_i)$.

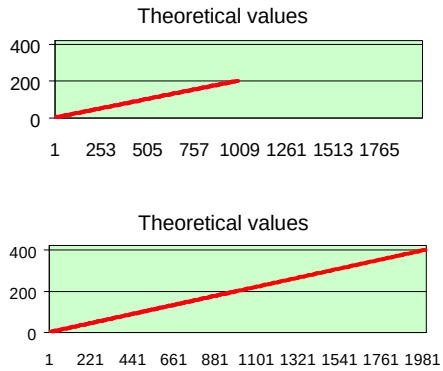
The simulation procedure take the following form :

- choice of a starting point x_t (for example, the initial individual in the LD matrix);
- determination of the winning $\overline{x}_i$ using $\underset{i}{\text{ArgMin}} \| x_t - \overline{x}_i \|$);
- picking at random an $\overline{y}_j$ according to the conditional distribution $P(\overline{y}_j | \overline{x}_i)$;
- computation of x_{t+1} by $x_t + \overline{y}_j$;
- iteration of the procedure for simulating the dynamics of the process on a specific time horizon.

For stochastic processes, the procedure will be iterated and the results will be averaged.

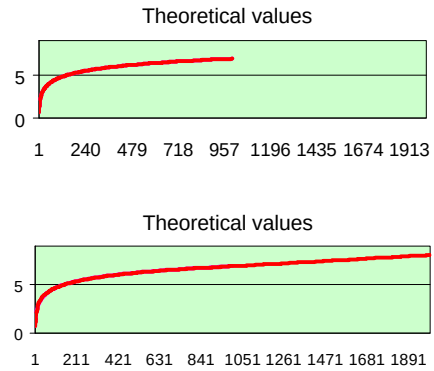
4. Applications

Three applications on deterministic processes are presented. For each one, a theoretical historic data set is formed by 2.000 individuals. Fig. 1 present the obtained results in the linear case, fig. 2 for a logarithmic function, fig. 3 for a sinusoidal one. The application to a Gaussian white noise is presented at fig. 4. The results seem self-speaking. The parameters used in each simulation are described in tab. 1.



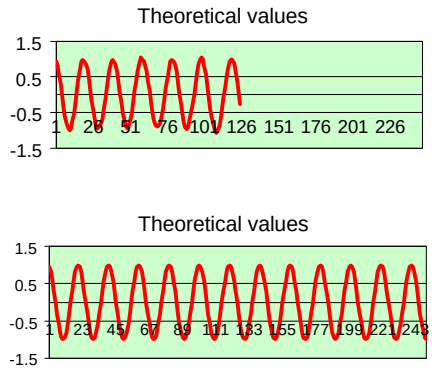
Correlation coefficient 1

Fig. 1. : Linear process reproduction and extrapolation



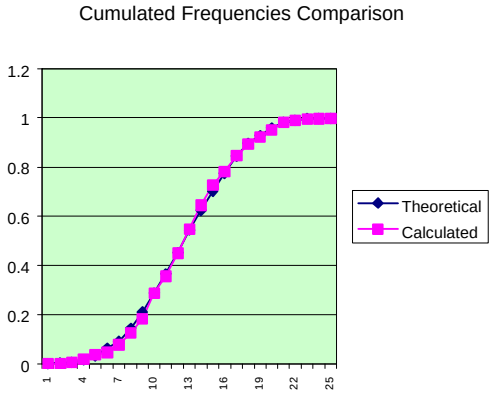
Correlation coefficient 0.999759

Fig. 2. : Logarithmic process reproduction and extrapolation



Correlation coefficient 0.964572

Fig. 3. : Sinusoïdal process reproduction and extrapolation



	Theoretical	Calculated
Mean	0	0.03054305
Stdev	1	0.995597383

Fig. 4. : Normal distribution reproduction

	Lin. process	Log. process	Sin. process	White noise
Equation	$y_t = 2 + (0.2 \times t)$	$y_t = \ln(t)$	$y_t = \sin(t)$	N(0,1)
t	t = 1 … 2000	t = 1 … 2000	$t = (\frac{p}{5}) \ \ldots \ (2000\ \frac{p}{5})$	
# clusters for LD	10	30	30	10
# clusters for P_i	10	10	10	10
λ	1	5	5	1
τ	1	1	1	1

Tab. 1 : parameters used for simulated processes.

The application of the procedure on a real data set (interest rate structures evolution over time) has been presented in [4]. We just present here at fig. 5 the ability of the procedure to reproduce the Cox, Ingersol & Ross [6] dynamic of short rate interest rate process, described by the following equation :

$$dr(t) = (a - br(t))dt + \acute{o}\sqrt{r(t)}dz(t)$$

where r(t) is the interest rate at maturity t, σ is the volatility of the process and z(t) is a standard Wiener process.

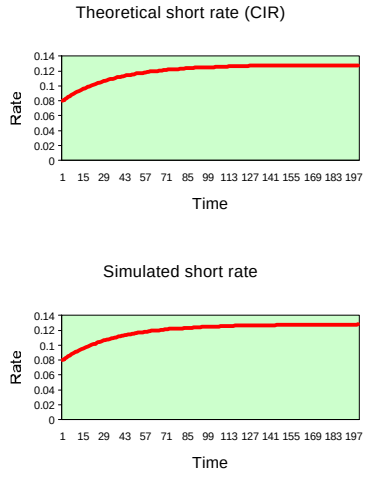


Fig. 5. :Theoretical short Rates produced by CIR Model / Simulated short rate on 200 periods

5. Bibliography

1. Cottrell M., Fort J.C., Etude d'un algorithme d'auto-organisation, *Annales de l'Institut Poincaré*, Vol. 23, 1, p. 1-20, 1987.
2. Cottrell M., Fort J.C., Pagès G., Two or three things that we know about the Kohonen algorithm, in *Proc of ESANN*, M. Verleysen ED., D Facto, Bruxelles, 1994.
3. Cottrell M., de Bodt E., Verleysen M., Kohonen maps versus vector quantization for data analysis, in *Proc of ESANN*, M. Verleysen ED., D Facto, Bruxelles, 1997.
4. Cottrell M., de Bodt E., Grégoire Ph., Simulating Interest Rate Structure Evolution on a Long Term Horizon : A Kohonen Map Application, *Proceedings of Neural Networks in The Capital Markets*, Californian Institute of Technology, World Scientific Ed., Passadena, 1996.
5. Cottrell M., Fort J.C., Pagès G., Theoretical aspects of the SOM algorithm, *Proc. of WSOM'97*, Helsinki, 1997.
6. Cox J.C., Ingersoll J.E., Ross S.A., *A Theory of the Term Structure of Interest Rates*, Econometrica 53, p. 385-407, 1985.
7. Kohonen T., Self-Organizing Maps, *Springer Series in Information Sciences*, Vol 30, Springer, Berlin, 1995.
8. Pagès G., Voronoï tesselation, Space quantization algorithms and numerical integration, *Proc. ESANN'93*, M.Verleysen Ed., Editions D Facto, Bruxelles, 221-228, 1993.
9. Walter J., Ritter H., Schulten K., Non-linear prediction with self-organizing maps, *Proc. of IJCNN*, San Diego, CA, 589-594, July 1990.