Change Detection in Nonstationary Time Series in Linear Regression Framework *

Theodor D. Popescu
National Institute for Research and Development in Informatics
8-10 Maresal Alex. Averescu Avenue, 71316 Bucharest, Romania
Email: pope@u3.ici.ro

The problem of segmenting the homogenous parts of a time series, or detecting abrupt changes in it is a key point that frequently arises in various application areas where modeling and processing of nonstationary time series is required. Many social and economic sciences have their own needs for time series analysis, which are closely related to segmenting and change detection in data.

The analysis of the behaviour of such real data reveals the most of the changes that occur are either changes in the mean level, or changes in spectral characteristics. The problem of segmentation between "homogenous" parts of the data (or detection of changes in the data) arises more or less explicitly. Actually, two main types of problems can be distinguished:

1. Segmentation of the data, the true model of which is not known, and where the model used for change or jump detection is simply a tool to locate the boundaries.

2. Segmentation of the data which are approximately represented by a large amount of models: the analysis is then of an artificial intelligence type, the changes may be not really abrupt.

The proposed problem formulation assumes off-line or batch-wise data processing, although the solution is sequential in data and an on-line data processing can be used. The segmentation model is the simplest possible extension of linear regression models to data with abruptly changing properties, or piece-wise linearizations of non-linear models. It is assumed that the data can be described by one linear regression within each segment with distinct parameter vector and noise variance.

The main goal of this project was to give a unified framework for the design and performance evaluation of some algorithms and methods for solving change detection problem in time series with application in econometrics. The following objectives have been taken into account:

1. To establish a methodological approach to deal with change detection in time series with application in the field of economics.

*This research was supported by a grant from the CERGE-EI Foundation under a program of the Global Development Network. Additional funds for grantees in the Balkan countries have been provided by the Austrian Government through WIIW, Vienna. All opinions expressed are those of the author(s) and have not been endorsed by CERGE-EI, WIIW, or the GDN

2. To evaluate the performances of some algorithms and methods for change detection in time series, presented in the literature, and to develop new methods and algorithms.

3. To design an integrated software support, implementing the best methods and algorithms for change detection in time series.

4. To prove the implemented methods and algorithms on case studies in the field of economics.

The Technical Report representing the first phase of the grant (Popescu 2002a) had as subject some of the best algorithms and methods for change detection in nonstationary time series. Also, some numerical results obtained by simulation have been presented, to rank the methods under consideration.

The limitations of some classical approaches, based upon the innovations of one AR model, have been outlined, and some new algorithms have been derived. The underlying idea of these algorithms is the use of two AR models: the first one is a global (or long term) model, which is estimated recursively, the second one is a local (or short term) one updated via a sliding block. Changes are detected when a suitable distance between these two models is too high. In a first step three "distance" measures are considered: cepstral distance, log-likelihood ratio and divergence between conditional probability laws. Other algorithms based on quadratic forms are also discussed. Hinkley's stopping-time may be coupled to these statistics in order to give a better estimate of the change time. All the algorithms have been compared via simulation study and has been investigated their robustness to the assumption of AR data.

Because the change detection in AR and ARMA problem is sometimes reduced to change detection in the mean for a sequence of independent observations governed by a Gaussian law $N(\mu, \sigma^2)$ before change and $N(\mu, \sigma^2)$ after the change, the main concepts and tools ("Filters Derivatives" algorithms, Shiryaev's algorithm and Hinkley's algorithm) which are used for online detection of model changes, in the simplest case, namely jumps in data mean, make the object of the study.

A problem of special interest discussed is change detection in a subset of model parameters, while the complementary subset of model parameters are completely unknown. A statistics for a global test (change in the AR part) is derived and used in change detection in modal characteristics of a single degree of freedom oscillator.

The problem of change detection in ARX and ARMAX models is solved in time and frequency domain. Several test indexes are derived, those developed in frequency domain being known by their robustness properties concerning system undermodeling and variations in experimental conditions. The algorithms are investigated by simulation for a single degree of freedom oscillator.

The Technical Report representing the second phase of the grant (Popescu 2002b) is dedicated to a software support implementing the best methods and algorithms for change detection in nonstationary time series.

First, an overview on the CHANGE program package for detection of changes, implemented on personal computers compatible IBM/PC is given. Then, some experimental results, obtained via simulation, for the test statistics, when the change detection methods based on quadratic forms are used are presented. Also, the robustness of the methods, as to the assumption of autoregressive data and to the model structure is discussed. Finally, the results obtained in segmentation of the following nonstationary financial and economic time series:

- US bond yield daily 1 April - 29 December 1989.
- UK bond yield daily 1 April - 29 December 1989.
- West Germany bond yield daily 1 April - 29 December 1989.
- Japan bond yield daily 1 April - 29 December 1989.
- 1 month - tbill monthly 30.01.1926 - 30.12.1996.
- US treasury bill 2nd market - middle rate, daily 11.06.1986 - 1.12.1995.

using the algorithms based on "distance" measure, quadratic forms and Kitagawa-Akaike method, are presented.

The significance of the research can be considered from two points of view:
From methodological point of view:

- To establish a unified and integrated approach for change detection in time series to be used in economics.
- To promote advances solutions (methods and algorithms) to problems in the field of analysis of economical processes.

From practical point of view:

- To propose a set of recommendations, based on the performance evaluation of the methods and on the case studies.
- To build an integrated software, implementing the best methods and algorithms for change detection problem solving in nonstationary time series analysis.

In conclusion, concerning the problem making the object of this project, we can mention the following remarks:

- Although the problem of change detection reached the maturity, there is a gap between theory and practice.
- The effort is now directed to robust change detection and diagnosis methods using reduced order models and adequate distance measures.
- These methods can not be reduced to repeated identification. Out purpose isn't to determine a good model, we use the model only like a tool in change detection schemes. Good and precise models offer high performance in change detection schemes, but also biased parametric models can be used for change detection and isolation. This bias decreases, but does not annihilate the performance of the detection procedure.

In our opinion, a coherent methodology is now available to the designer, together with the corresponding set of tools, which enables him to solve a large variety of change detection problem in dynamical systems. The general opinion of the scientific community with preocupations in this field is that there is a gap between theory and practice and that the model based methods have many more possibilities in the real practical problems than they so far have proved to have.

The topics of change detection are of increasing practical importance and therefore theoretical as well as applied research is a challenge for the future.

References

Popescu, Th., Change Detection in Nonstationary Time Series in Linear Regression Framework. Algorithms and Methods, Technical Report - Grant CERGE-EI Foundation, 2002a

Popescu, Th., Change Detection in Nonstationary Time Series in Linear Regression Framework. Software Support and Applications, Technical Report - Grant CERGE-EI Foundation, 2002b