1. Introduction

We consider the problem of modeling a non-stationary time series by segmenting the series into blocks of different autoregressive (AR) processes. The number of break points, denoted by $m$, as well as their location, and the order of the respective AR models are assumed to be unknown. We propose an automatic procedure for obtaining such an optimal partition called Auto-PARM for Automatic Piecewise AutoRegressive Modeling.

1.1 Piecewise Autoregressive processes

Setup: there exist $m$ and $t_{\gamma_j} = 1 < t_{\gamma_1} < \ldots < t_{\gamma_m} = n + 1$ ($n =$ sample size) such that

$$ Y = y_{\gamma_1} + \theta_{\gamma_1} y_{\gamma_1-1} + \cdots + \theta_{\gamma_j} y_{\gamma_j-1} + \epsilon_{\gamma_j}, \quad \text{if} \quad t_{\gamma_j} < t < t_{\gamma_{j+1}},$$

where $\{\epsilon_{\gamma_j}\}$ is IID $N(0,1)$.

Goal: Estimate $m$ — number of segments $t_{\gamma_j}$ — location of $j$th break point $\gamma_j$ — level in $j$th epoch $\theta_{\gamma_j}$ — order of AR process in $j$th epoch $\epsilon_{\gamma_j}$ — AR coefficients in $j$th epoch $\sigma_{\gamma_j}$ — scale in $j$th epoch

1.2 Motivation for using piecewise AR models

Piecewise AR is a special case of a piecewise stationary process (see Adak 1998) $Y_{t,n} = \sum_{j=1}^{\infty} \gamma_j I_{\{t \leq t_j\}}(t/n)$, where $\{Y_{t,n}\} \sim \{\epsilon_{\gamma_j}\}$ is a sequence of stationary processes. It is argued in Onobah et al. (2001) that if $\{Y_{t,n}\}$ is a locally stationary process (in the sense of Dahlhaus), then there exist a piecewise stationary $\{Y_{t,n}\}$ and a sequence $m_{\gamma_j}$,

$m_{\gamma_j} \rightarrow \infty$ and $m_{\gamma_j} / n \rightarrow 0$ as $n \rightarrow \infty$,

that approximates $\{Y_{t,n}\}$ (in average mean square).

Roughly speaking, $\{Y_{t,n}\}$ is a locally stationary process if it has a time-varying spectrum that is approximately $(1/t)F(t)$, where $d(F)$ is a continuous function in $n$.

1.3 Basics of the Genetic Algorithm (GA)

The GA is an optimization algorithm that mimics natural evolution.

- Start with an initial set of chromosomes, or population, of possible solutions to the optimization problem.
- Parent chromosomes are randomly selected (proportional to the rank of their objective function values), and produce offspring using crossover or mutation operations.
- After a sufficient number of offspring are produced to form a second generation, the process then restarts to produce a third generation.
- Based on Darwin’s theory of natural selection, the solution should produce future generations that give a smaller (or larger) objective function.

2. Implementation of GA

A chromosome consists of $n$ genes, each taking the value of -1 (no break) or 0 (order of AR process). Use natural selection to find a near optimal solution. An element $F_{\theta'}$ is mapped with a chromosome $c$ by

$$(m, (\gamma_1, \gamma_2, \ldots)) \rightarrow c = (\delta_1, \delta_2, \ldots),$$

where $F_{\theta'}$ is a continuous function in $n$.

2.2 The MDL applied to piecewise AR models

MDL = class of piecewise AR models for $y = (y_1, \ldots, y_n)$

$L_{\theta_j}(y) =$ code length of $y$ relative to $F \in M$.

Best fitting MDL model is minimizer of

$$MDL(m, (\gamma_1, \gamma_2, \ldots), (\gamma_3, \gamma_4, \ldots)) = \log n + \log \sum_{j=1}^{m} \log p_j + \sum_{j=1}^{m} \sum_{i=1}^{n} \log \sigma_{\gamma_j}^2,$$

where $\sigma_{\gamma_j}$ is the length of the $j$th segment and $\gamma_j$ is the Yule-Walker estimate of $\sigma_j$ in the $j$th segment.

2.3 Consistency

Assume there exist true values $m$ and $\{\gamma_j\}$ with $0 < \gamma_1 < \gamma_2 < \ldots < \gamma_m < 1$ with $t_{\gamma_j} = [\gamma_j n], j = 1, 2, \ldots, m$.

Theorem: For the piecewise AR model, if the number of breakpoints $m$ is known, then

$$\hat{\gamma}_j \rightarrow \gamma_j, \quad \text{a.s.}, \quad j = 1, 2, \ldots, m.$$

2.4 Slowly varying AR(2) model

$$Y = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \epsilon_{t}, \quad \text{if} \quad 1 \leq t \leq 1024$$

where $\{\epsilon_{t}\} \sim N(0,1)$.

4.1 Piecewise stationary with dyadic structure

$$y = \theta_0^{\ast} + \theta_1^{\ast} Y_{t-1}, \quad \text{if} \quad 1 \leq t \leq 512$$

$$y = \theta_2^{\ast} + \theta_3^{\ast} Y_{t-1} + \theta_4^{\ast} Y_{t-2} + \epsilon_{t}, \quad \text{if} \quad 513 \leq t \leq 1024$$

2.5 Implementation using Minimum Description Length (MDL)

The problem behind MDL is to choose the model which maximizes the compression of the data or, equivalently, select the model that minimizes the code length of the data (i.e., amount of memory required to encode the data).

3.2 Model selection using Minimum Description Length (MDL)

For example, $c = (1, -1, -1, -1, 0, -1, -1, -1, -1, 1, -1, 1, 1, -1, -1, -1)$

corresponds to

AR(2), $t = 5$; AR(3), $t = 11$; AR(0), $t = 11$; AR(3), $t = 15$; AR(2), $t = 15$.

5. Conclusions

- Introduced Auto-PARM (an automatic procedure for segmenting a time series into piecewise AR models).
- Model selection based on MDL (minimum description length principle).
- A genetic algorithm was used to find a near optimal solution to the model selection problem based on MDL.
- Auto-PARM works well for both detecting segments and for estimating time-varying spectra.

6. References