Summer Research Project

2005 – 2006

The Distributors of Change Points in Long Memory Processes

Guan Yu Zheng

Department of Mathematics and Statistics
University of Canterbury
The distributions of change points in long memory processes

GuanYu Zheng

Supervisors: Marco Reale and Bill Rea

Department of Mathematics and Statistics

University of Canterbury

New Zealand

February 9, 2007

Abstract

In this paper we present the properties of empirical distributions of different statistics (e.g. standard deviation and number of breaks) related to the presence of structural breaks in simulated Fractional Gaussian Noise series with various Hurst parameters. Structural Breaks are detected with Atoretical Regression Trees, a structural break identification method. The simulation results were applied to four case studies to check whether a Regime Switching or Fractional Gaussian Noise model is more adequate.

Keywords: Long-range dependence, strong dependence, global dependence, Hurst phenomena.

Introduction

The concept of long memory was introduced by Mandelbrot in 1965. It also is known as long-range dependency, strong dependence or Hurst phenomenon. Both of statistic and economic literatures have made great effort to develop this field of study. Up to
now, the presence of long memory time series is becoming more and more important in new areas like finance and network communications as well as hydrology, geophysics and climatology.

The intention of this paper is to provide a simple approach to test whether a given dataset is better described by a long memory model or a regime switching model.

The next sections of this paper are organized as follows: section (2) gives a brief definition of long memory process and self-similarity. Section (3) presents overview of Fractional Integration (FI), Fractional Gaussian Noise (FGN) and Regime Switching (RS). Section (4) describes the distributional properties of structural breaks for different Hurst coefficients. Section (5) presents the results of our investigation of self-similarity properties of four real datasets. Section (6) presents the conclusions.

2 Definition of Long memory and Self-similarity

In this section we introduce the concept of long memory and self-similarity. Then, an illustration is provided of the equivalence of self-similarity and long memory.

2.1 Long memory

There are several ways of defining the long memory of a discrete time series formulated both in time and frequency domain. A simple definition proposed by Baillie [2] was given in econometric literature and illustrated below.

Definition 1 The process possesses long memory if

\[ \lim_{n \to \infty} \sum_{j=-n}^{n} |\rho(j)| \to \infty \]

Equivalently, the spectral density \( f(\omega) \) of the data will be unbounded at low frequencies.
Figure 1: Time series plot, autocorrelation and raw periodogram with tree ring width in Campito Mountain (Left) and a simulated ARMA (1,1) with $p = 0.5$ and $q = 0.5$.

Silverberg and Verspagen [3] pointed out that the decay rate of the Autocorrelation Function (ACF) in a short memory process (ARMA) is about exponential with a zero spectral density at the origin. However, the ACF of a long memory decays at a hyperbolical rate which is much slower than the exponential rate found for stationary ARMA processes. Furthermore, the power of spectrum is completely dominated by the low frequency component.

And how about trying to first difference the raw dataset, and what changes can we find? The whole time series is smoothed, and appear more stationary. Even though the ACF decays very quickly, its values still stay at some significant level. Furthermore, the power in the origin of periodogram is perished. In other words, there is no power in the origin. There are clear signs that the data is over-differenced.
Figure 2. ACF (above) and Periodogram (Below) of first differenced Campito Mountain.

The Autoregressive Fractional Integrated Moving Average (ARFIMA) model was derived to overcome this type of problem.

**Formula 1** The ARFIMA (p,d,q) model is expressed as

\[ \phi(B) \Delta^d \chi = \theta(B) \epsilon \]

Where \( (\epsilon) \) is a White Noise with zero mean, and \( \Delta^d = (1 - B)^d \), where \( d \in (0,0.5) \), is a difference operator.

In this paper, the simplest ARFIMA model, which is the fractional integrated noise (FI) or ARFIMA (0,d,0), is considered
The FI(d) model is expressed as
\[ \Delta^d x_t = \epsilon_t \]

Its counter-part for continuous time series is the Fractional Gaussian Noise (FGN). We would explain it later. The relationship between \( d \) and the Hurst coefficient \( H \) is: \( H = 0.5 + d \).

2.2 Self-similarity

The concept of self-similar process was first introduced by Kolmogorov and its definition as presented by Mandelbrot and van Ness [4] is:

**Definition 2** a stochastic process \((X(t), t \geq 0)\) is self-similar with Index \( H > 0 \), if
\[
\left( X(ct), t \geq 0 \right) \quad \text{and} \quad \left( c^H X(t), t \geq 0 \right)
\]
have the same finite dimensional distribution. \( H \) is called Hurst parameter or self-similar parameter with values between 0 and 1.

In other words, the stochastic points have the same statistical properties at different periods of time in time series data.

2.3 Connection between long memory and self-similarity

The auto-correlation function can be expressed as
\[
\rho(k) = \frac{1}{2} \left[ (k+1)^{2H} - 2k^{2H} + (k-1)^{2H} \right],
\]
in term of self-similarity. When \( H = 0.5 \), the process is a White Noise. Interestingly, the correlations decay at very slow rate and are not summable for \( 0.5 < H < 1 \). This implies \( \sum_{k=-\infty}^{\infty} \rho(k) \rightarrow \infty \); therefore, a self-similar process with \( \frac{1}{2} \leq H < 1 \) can be considered a long memory process.

3 Model

This paper considers two alternative models the Regime Switching (RS) vs. Fractional Integration (FI) or Fractional Gaussian Noise (FGN), and they are applied to four case studies to help identifying which model is more appropriate to describe them.

The FGN can be defined from definition 2 and below from Mandelbrot and van Ness [4].
Definition 3: a stochastic process \((X(t), t \geq 0)\) has stationary increment, if

\[ X(t + c) - X(c) \equiv_d X(t) - X(0) \]

The time series plots were given below for illustration. Another point we should notice is that the FGN process is self-similar if the \(H\) parameter is constant through the series.

![Time series plots](image)

Figure 3: Four different Fractional Gaussian Noise time series plot.

The alternative model is a Regime Switching or structural break model presented by Ohanissian, Russell and Tsay[5], Granger and Terasvirta [6]. This model contains stationary sub-time series (stationary ARMA models) with probabilistic changes on state levels. Its formulation was presented by Chen and Tiao [7] is as follows.
Formula 3: a (discrete) time series \( y_t \) if

\[
y_t = \mu_t + x_t \tag{1}
\]

\[
\mu_t = p_t \eta_t + \mu_{t-1} \tag{2}
\]

Where \( (x_t, t \geq 0) \) is local-stationary ARMA model, \((\mu_t, t > 0) \in N(0, c \sigma^2)\), and \( p_t \) is binary variable with \( \text{Prob}(p_t = 1) = \alpha \) and \( \text{Prob}(p_t = 0) = 1 - \alpha \).

In Regime Switching, the mean levels change over time according to the probability of \( p_t \). During the changes or breaks, the series are stationary and have different statistical properties. Hence, Regime Switching is not self-similar.

Some authors, like Russell, Ohanissian and Tsay [5], claimed that RS can be distinguished from a long memory process. But the two alternative models are almost observationally equivalent. Hence the two models are difficult to tell apart.

In section 5, we would provide some distribution tests for the four datasets and we shall try to tell whether a FGN or RS is more appropriate.

4 Properties of simulated Fractional Gaussian Noises

In this part, nine FGN series with different Hurst parameters (H=0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9 and 0.95) were simulated 1000 times with the statistical package fseries[8] in R[9], and were broken into regimes by ART (Cappelli and Reale[10]). The length of each series consisted of 5405 sequential points which is the same as example of the Capito Mountain. Then the empirical distributions of some statistics were derived from evaluating the statistical properties on these regimes. These statistics were number of breaks, regime length, mean level, standard deviation, skewness and kurtosis. Another R package MASS was used to estimate the distribution parameters by the Maximum Likelihood approach.

4.1 Number of breaks

The number of breaks is an integer value, and its distribution is discrete. The Poisson distribution gave the best approximation and was applied. The Normal distribution was used as an alternative.

The shape of the distribution was symmetric when H was larger than 0.9. Then, it turns to be right-skewed as H became smaller. Hence, we found that the Poisson distribution fitted well the distribution for large H values (H≥0.8).
Table 1: Expected number of breaks in FGNs.

<table>
<thead>
<tr>
<th>H value</th>
<th>Expected number of breaks</th>
<th>H value</th>
<th>Expected number of breaks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>8.29</td>
<td>0.7</td>
<td>0.75</td>
</tr>
<tr>
<td>0.9</td>
<td>6.82</td>
<td>0.65</td>
<td>0.15</td>
</tr>
<tr>
<td>0.85</td>
<td>5.01</td>
<td>0.6</td>
<td>0.0015</td>
</tr>
<tr>
<td>0.8</td>
<td>3.19</td>
<td>0.55</td>
<td>0</td>
</tr>
<tr>
<td>0.75</td>
<td>1.76</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The distribution changed to exponential-like shape if $H$ was less than 0.8. As the number of breaks tend to zero, but this also meant that for $0.5 \leq H < 0.8$ it was not a problem to distinguish between long memory and RS. Another point we noticed is that the expectation of number of breaks was shrunk as $H$ was lessened. That implies that the FGN processes became “stationary” as no breaks were detected.
4.2 Regime length

The distribution of original regime length is highly right-skewed and gamma distributed. It did not present well-behaved features. Natural logarithm transformation was used and gave some interesting results.

After log-transformation, distribution is less skewed, and it is quite symmetric at $0.9 \leq H < 1$. Gamma and Normal distribution were used. The Gamma distribution probably was the better one to capture the changes in the empirical distribution. However, distribution was atypical at small H value ($0.5 \leq H < 0.8$).
The expected value of the regime length became larger and larger as $H$ became smaller and smaller. This is consistent with the findings of the previous results as the number of breaks tends to zero. At $0.5 \leq H \leq 0.65$, the regime length was very likely to be equal to the length of the series.

<table>
<thead>
<tr>
<th>$H$ value</th>
<th>Expectation of log regime length</th>
<th>$H$ value</th>
<th>Expectation of log regime length</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>6.46 (637.61)</td>
<td>0.7</td>
<td>8.23 (3763.93)</td>
</tr>
<tr>
<td>0.9</td>
<td>6.65 (771.26)</td>
<td>0.65</td>
<td>8.52 (5004)</td>
</tr>
<tr>
<td>0.85</td>
<td>6.95 (1042.03)</td>
<td>0.6</td>
<td>8.59 (5388.83)</td>
</tr>
<tr>
<td>0.8</td>
<td>7.37 (1584.58)</td>
<td>0.55</td>
<td>8.6 (5405)</td>
</tr>
<tr>
<td>0.75</td>
<td>7.81 (2470.29)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Expectations of log regime length and original regime length.
4.3 Mean levels
Differently from the previous two statistics, the empirical distribution of the mean always appears Normal. The centre was zero, and the sample standard deviation became smaller as H moved from 0.95 to 0.55. The Logistic distribution was used as well. It might be a reasonable alternative.

![Dist of mean with H=0.95](image1)

![Dist of mean with H=0.85](image2)

![Dist of mean with H=0.8](image3)

![Dist of mean with H=0.7](image4)

![Dist of mean with H=0.65](image5)

![Dist of mean with H=0.55](image6)

Figure 6: Empirical distributions of mean. Normal (solid) and Logistic (dash)

4.4 Skewness
In term of the skewness, the empirical distribution showed symmetry through all H values. Its centre shifts from 0.8 to 1, and the density of point at 1 becomes larger and larger as H becomes smaller. In other words, the skewness is more stable where H is small. When H=0.55, the skewness was 1. Additionally, densities along the two-sides decayed very quickly.

Because of these features, Logistic and normal distribution worked well where $0.8 \leq H < 1$. The Logistic just did a little better than the Normal for this case.
The rest might be estimated by double-exponential distributions.

Figure 7: Empirical distributions of standard deviation. Normal (Solid) and Logistic (Dash)

4.5 Kurtosis
Similar to the section 4.4, symmetry and diminished dispersion were apparent features in these distributions through all Hurst coefficients. The only difference was that the centre of the distribution was approximately zero.

4.6 comments
In this section, the distributions of statistics for FGN’s with different H coefficients were generally identified at large H values between 0.8 and 1 whereas further studies are needed for $0.5 \leq H < 0.8$. 
5. Discussion

In this section, we applied the results of section 4 to four case studies, and drew some useful information to tell whether long memory with stationary increment or regime switching is more appropriate. The description of all datasets was given by Rea et al [11].

In the next paragraphs, we present the distributions of number of breaks and log regime length as they appeared particular effective in identifying between the two alternative models. Poisson and gamma distribution were respectively applied. Moreover, Figure 9-12 provided empirical distributions on each case, and Figure 13-16 presented time series and its structural breaks by ART.

The results for the other statistics are available on requests from the author.
5.1 Nile Minima

The estimated Hurst parameter for the Nile Minima was 0.837 given by the Whittle estimator in fseries. ART returned 10 breaks i.e. 11 regimes in this dataset. The expected number of breaks in the simulated series for the corresponding H is 11.18. The probability of having 10 breaks is 11.7 percent which was almost the highest through all the densities. Hence, this can be considered a long memory process at 0.95 confidence level.

In term of regime length, the simulation showed the expected log regime length was 3.64 (38.09). And, log regime length from Nile Minima was 4.33 (76.25). That supported the hypothesis that the Nile River followed a FGN process.

![Hist of simulated number of break at FGN H=0.837](image)

![Hist of simulated regime length at FGN H=0.837](image)

Figure 9: Empirical distribution of number of breaks and log regime length in Nile yearly minimum water levels 662 to 1284AD.

5.2 Camptio Mountain

ART found 12 break points in Campito Mountain. However, the expected break point was 5.97 given FGN with H=0.876. The chance of having 12 was smaller than 1 percent.
Compared to 6.36 (578) from simulated data, the sample mean of log regime length was slightly smaller (6.2, (491.27)) with a little error.

The Campito Mountain data is considered a class example of a FI process. However the result based on our simulations suggested the regime switching was more suited thus supporting the results of Rea et al [11].

Hist of number of breaks with FGN $H=0.876$

Hist of log regime length of Campito with $H=0.876$

Figure 10: Empirical distribution of number of breaks and log regime length in tree ring width in Capito Mountain 3435BC to 1969AD.

5.3 Shihua Cave
The break number was 14 in this dataset. Whereas the expectation of the number of breaks was only 5.79. The probability of having 14 breaks was 0.0017. That implies the FGN model poorly performed against the RS.

Similarly, the distribution of the long regime length is far from the gamma distribution. Sample average of log regime length was 7.25 (1416), and expected length from
gamma was 5.65 (285.08). Obviously FGN and Self-similarity were inconsistent with the data.

\[ \text{Hist of simulate Shihua cave with FGN } H=0.83 \]

\[ \text{Hist of log regime length with FGN } H=0.83 \]

Figure 11: Empirical distribution of number of breaks and log regime length in thickness of annual layers of a stalagmite in Shihua Cave 665BC to 1985AD.

5.4 Elk lake
Like the previous two cases, the empirical Poisson and gamma distribution presented an apparent contrast with the observed results. The data was not FGN.

The data was given 8 break points and 9 regimes. For a self-similar long memory process of this distribution we have approximately 4.15. The probability of having 8 breaks was just 3 percent. In addition, the sampled log regime length was 8.58 (5320) which was much larger than 7.33 (1526.85) from model.

Bill et al [11] also introduced infrequent changes on mean through time scales. That strongly suggested Regime switching would describe to data better.
6. Conclusion

We have proposed a data driven parametric procedure to distinguish between long memory model and RS. The technique is to generate FGNs, introduce parametric distributions for each quantity by breaking series using ART, and compare to actual data. Four data sets are considered and the long memory behavior tests.

We found that distributions were easy to derive when Hurst parameter is high enough (usually larger than 0.8). When Hurst was lower than that point, some other tools needed to form theoretical distribution model. But given the low likelihood of having breaks for $0.5 \leq H < 0.8$, this study provides a useful tool to distinguish between long memory and RS.

Three of the four examples exhibit regime switching property and non-self-similarity
with clear results. The results here obtained can be exploited to construct a test for self-similarity vs Regime Switching.

Figure 13: Time series of Nile river yearly minimum water levels and its ART.
Figure 14: Time series of Capito mountain and its ART.

Figure 15: Time series of thickness in Shihua Cave and its ART
Figure 16: Time series of thickness of Elk lake varve sequence and its ART.

References


