Proceedings of the 28th Annual IEEE Conference on Local Computer Networks, pp. 307-308, Bonn, Germany, October 2003

ESTIMATION OF THE LONG-RANGE DEPENDENCE PARAMETER OF FRACTIONAL ARIMA PROCESSES

Houssain Kettani
Department of Computer Science
Jackson State University
Jackson, MS 39217
kettani@uwalumni.com

John A. Gubner
Department of Electrical and Computer Engineering
University of Wisconsin
Madison, WI 53706
gubner@engr.wisc.edu

1. Introduction

The most well-known models of long-range dependent processes are fractional Gaussian noise [7] (thus second-order self-similarity) and fractional ARIMA [3, 4]. Each of these models has a corresponding long-range dependence parameter. Since the value of the parameter indicates the intensity of this dependence structure, it is important to have a better tool to estimate it. Such an estimator should not be biased and the confidence intervals should be as small as possible. Moreover, if we would like to estimate the parameter on-line, then the estimation tool should be as fast as possible.

Several methods for estimating long-range dependence parameters have been proposed; see [2] for details. By far, the wavelet method is the most widely used. When a process is assumed to be second-order self-similar, [6] introduced a new method that uses the structure of the covariance function to estimate the Hurst parameter. The method was shown to be much faster and yield smaller confidence intervals than the wavelet method. In this paper, we consider the case when the process is assumed to be fractional ARIMA and show that the new method still possesses the aforementioned qualities.

2. Preliminaries

Let X_i denote the number of bits, bytes, or packets seen during the ith interval. We say that X_i is *second-order stationary* if its mean $E(X_i)$ does not depend on i and if the autocovariance function $E[(X_i-E(X_i))(X_j-E(X_j))]$ depends on i and j only through their difference $k=i-j$, in which case we write $\gamma(k)=E[(X_{i+k}-E(X_{i+k}))(X_i-E(X_i))]$. We then put $\sigma^2=\gamma(0)=E[(X_i-E(X_i))^2]$, and $\rho(k)=\frac{\gamma(k)}{\sigma^2}$, to denote the variance and autocorrelation function of the process X_i, respectively.

The fractional ARIMA(p,d,q) process proposed in $[3,4]$ is an extension of the ARIMA(p,d,q) in the sense that d is allowed to take any real value in the interval $(-\frac{1}{2},\frac{1}{2})$. Any fractional ARIMA$(p,d,q)$ process can be expressed in terms of the *standard* fractional ARIMA$(0,d,0)$. The autocorrelation function of the latter is given by

$$\rho(k)=\prod_{i=1}^{k}\frac{k-i+d}{k-i+1-d}. \tag{1}$$

3. Main Result

Given observed data $X_1,\dots,X_n$, let $\hat{\mu}_n = n^{-1}\sum_{i=1}^{n}X_i$, $\hat{\gamma}_n(k)=\frac{1}{n}\sum_{i=1}^{n-k}(X_i-\hat{\mu}_n)(X_{i+k}-\hat{\mu}_n)$, $\hat{\sigma}_n^2=\hat{\gamma}_n(0)$, and

$$\hat{\rho}_n(k)=\frac{\hat{\gamma}_n(k)}{\hat{\sigma}_n^2}, \tag{2}$$

denote the *sample mean*, the *sample covariance*, the *sample variance*, and the *sample autocorrelation*, respectively.

Suppose that X_i is known to be fractional ARIMA$(0,d,0)$. Then from (1) we have $\rho(1)=d/(1-d)$. Thus, we propose

$$\hat{d}=\frac{\hat{\rho}_n(1)}{1+\hat{\rho}_n(1)}, \tag{3}$$

as an estimate of the difference parameter of the process X_i.

To assess the performance of the proposed estimate, we appeal to the following result due to Anderson [1] and Hosking [5].

Theorem: *Let X_i be a fractional ARIMA$(0,d,0)$ process. If the process is Gaussian, then for large sample size n, $\hat{\rho}_n(1)$ has mean*

$$\mu_n=\rho(1)-\frac{(1-2d)}{d(1-d)(1+2d)}\frac{\Gamma(1-d)}{\Gamma(d)}n^{2d-1},$$

and

1. *If $d \in (-\frac{1}{2}, \frac{1}{4})$, then $\hat{\rho}_n(1)$ is approximately N(μ_n, σ_n^2) with*

$$\sigma_n^2 = \frac{2}{n}\left[\frac{1-2d}{1-d}\right]^2, \tag{4}$$

2. *If $d = \frac{1}{4}$, then $\hat{\rho}_n(1)$ is approximately N(μ_n, σ_n^2) with*

$$\sigma_n^2 = \left[\frac{2(1-2d)\Gamma(1-d)}{(1-d)\Gamma(d)}\right]^2 \frac{\log n}{n} \tag{5}$$

3. *If $d \in (\frac{1}{4}, \frac{1}{2})$, then the limiting distribution of $\hat{\rho}_n(1)$ has mean μ_n and variance given by*

$$\sigma_n^2 = 2\left[\frac{(1-2d)\Gamma(1-d)}{(1-d)\Gamma(d)}\right]^2 K_2(d)n^{4d-2}, \tag{6}$$

where $K_2(d)$ is related to the variance of the modified Rosenblatt distribution and is given by

$$K_2(H) = \int_0^1 \int_0^1 g^2(x,y)dxdy,$$

where

$$g(x,y) = |x-y|^{2d-1} - \frac{1}{2d}(x^{2d}+y^{2d}+(1-x)^{2d}+(1-y)^{2d}) + \frac{1}{d(2d+1)}.$$

With σ_n^2 given by the theorem, we have $d_- \leq \hat{d}_n \leq d_+$, with 95% probability, where

$$d_\pm = \frac{\mu_n \pm 1.96\sigma_n}{1 + \mu_n \pm 1.96\sigma_n}. \tag{7}$$

Now, for known d, the 95% confidence interval of the estimate $\hat{d}_n$ is $[d_-, d_+]$. Next, let w_n denote the width of such intervals, i.e., $w_n = d_+ - d_- \cdot$ A log-log plot of w_n versus the number of samples n for different values of d resembles straightlines. Thus, the width w_n can be written as

$$w_n \approx an^{-b}, \tag{8}$$

where a and b are constants for fixed d value. The values of these constants are given in Table 1.

It is interesting to note that the width w_n is upper bounded by the w_n at the value $d = -0.40$. Hence in the case when d is not known, we choose the confidence interval centered around $\hat{d}_n$ with width

$$w_n = \frac{15}{\sqrt{n}}. \tag{9}$$

d	a	b
-0.40	14.88	0.50
-0.30	11.53	0.50
-0.20	9.32	0.50
-0.10	7.33	0.50
0.00	5.57	0.50
0.10	4.05	0.50
0.20	2.79	0.50
0.25	1.92	0.45
0.30	1.51	0.41
0.40	0.21	0.21

Table 1. The values of the constants a and b in (8) for different d values.

4. Summary of the Algorithm

In what follows, we present a summary of the new method:

- Let $X_1, X_2, \ldots, X_n$ be a realization of a fractional ARIMA$(0, d, 0)$ process,
- Compute $\hat{\rho}_n(1)$ as in (2),
- Compute $\hat{d}_n$ as in (3), which is the estimated difference parameter,
- The 95% confidence interval of d is centered around the estimate $\hat{d}_n$ with width as in (9).

We have compared our proposed method with the wavelet method using simulated fractional ARIMA data and real data. We found that our proposed method is much faster and yields smaller confidence intervals than the wavelet method. Similar conclusions were made for our similar method for fractional Gaussian noise data [6].

References

[1] T. W. Anderson (1958). *The Statistical Analysis of Time Series*. John Wiley & Sons, New York, New York.

[2] J. Beran (1994). *Statistics for Long-Memory Processes*. Chapman & Hall, New York, New York.

[3] C. W. J. Granger and R. Joyeux, (1980). An Introduction to Long-Memory Models and Fractional Differencing *Journal of Time Series Analysis*, 1:15–29.

[4] J. R. M. Hosking (1981). Fractional Differencing. *Biometrika*, 68:165–176.

[5] J. R. M. Hosking (1996). Asymptotic Distribution of the Sample Mean, Autocovariances, and Autocorrelations of Long-memory Time Series. *Journal of Econometrics*, 73:261–284.

[6] H. Kettani and J. A. Gubner (2002) A Novel Approach to the Estimation of the Hurst Parameter in Self-Similar Traffic. *Proceedings of the 27th Annual IEEE Conference on Local Computer Networks.*

[7] B. B. Mandelbrot (1983). *The Fractal Geometry of Nature*. W. H. Freeman and Co., New York, New York.