

COMPARISON OF DIFFERENCING PARAMETER ESTIMATION FROM NONSTATIONER ARFIMA MODEL BY GPH METHOD WITH COSINE TAPERING

Gumgum Darmawan

Lecturer of Statistics Department of Padjadjaran University;

gumstat_1973@yahoo.com,

ABSTRACT

This paper compares the accuracy of estimation of differencing parameter from nonstationary ARFIMA Data. For comparison purposes, the GPH (Geweke and Porter-Hudak) method is modified by three tapers, there are Cosine bell, Hanning and Hamming tapers. Accuracy among three methods are justified by Mean Square Error and Deviation Standard in two models, i.e ARFIMA(1,d,0) and ARFIMA(0,d,1). From simulation results, GPH method with Cosine Bell tapering shows a good performance in estimating the differencing parameter of ARFIMA(0,d,1) data. From ARFIMA(1,d,0) data, GPH method with Hanning tapering is the best of all methods

Keywords : ARFIMA , Periodogram, Taper

1. INTRODUCTION

Long range dependence or long memory means that observations far away each other are still strongly correlated. The correlation of a long-memory process decay slowly that is with a hyperbolic rate, not exponentially like for example ARMA-process (see figure.1)

The literature on ARFIMA processes has rapidly increased since early contribution by Granger and Joyeux (1980), Hosking (1981) and Geweke and Porter-Hudak (1983). This theory has been widely used in different fields such as meteorology, astronomy, hydrology and economics.

Geweke and Porter-Hudak (1983) presented a very important work on stationary long memory processes. Their paper gave rise to several other works, and presented a proof for the asymptotic distribution of the long memory parameter. These authors proposed an estimator of d as the ordinary least squares estimator of the slope parameter in a simple linear regression of the logarithm of the periodogram. Reisen (1994) proposed a modified form of the regression method, based on a smoothed version of the periodogram function. Robinson (1995) has made a mild modifications on Geweke and Porter-Hudak's estimator, dealt simultaneously with $d \in (-0.5, 0.0)$ and $d \in (0.0, 0.5)$. Hurvich and Deo (1999), among others, addressed the problem of selecting the number of frequencies necessary for estimating the differencing parameter in the stationary case. Fox and Taqqu (1986) considered an approximation method, whereas Sowell (1992) presented the exact maximum likelihood procedure for estimating the fractional parameter. These two papers considered the estimation procedures for the stationary case. Simulation studies comparing estimates of d may be found, for instance, in Bisaglia and Guegan (1998), Reisen and Lopes (1999).

Here, by simulation study, we compare accuracy of estimation the differencing parameter of nonstationary ARFIMA model by the spectral regression methods with cosine tapered periodogram. The first estimator is the modified periodogram function where the tapered data is obtained from the cosine-bell function, this estimator was used in the works by Hurvich and Ray (1995) and Velasco (1999a). As the second and

the third estimator, we consider Hanning and Hamming tapers respectively for estimating the differencing parameter of ARFIMA model. Autoregressive and Moving average parameters are estimated by maximum likelihood.

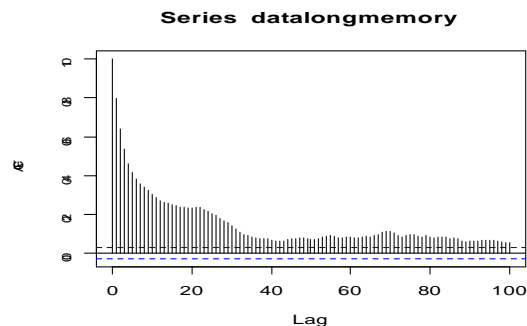


Figure 1. ACF plot of long memory processe

2. ARFIMA MODEL

A well known class of long memory models is the autoregressive fractionally integrated moving average (ARFIMA) process introduced by Granger and Joyeux (1980) and Hosking (1981).

An ARFIMA model (p,d,q) can be defined as follows:

$$\phi(B)(1-B)^d(z_{t-\mu})\theta(B)a_t \tag{1}$$

where

t = index of observation ($t = 1,2,\dots,T$)

d = the degree of differential parameter (real number)

μ = mean of observation

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, polynomial of AR(p)

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$, polynomial of MA(q)

$(1-B)^d$ = fractional differencing operator

a_t is IID($0, \sigma^2$).

3. GPH METHOD

There have been proposed in the literature many estimators for the fractional differencing parameter. We shall concentrate in estimators based upon the estimation of the spectral density function. The semi-parametric estimators describe below are obtained taking the logarithm of the spectral density. Estimation of d from ARFIMA (p,d,q) model as follows

a) Construct spectral function of ARFIMA (p,d,q) model

For the ARFIMA model given in equation (1), let $w_t = (1-B)^d z_t$, and let $f_w(\omega)$ and $f_z(\omega)$ be the spectral density function of $\{w_t\}$ and $\{z_t\}$, respectively. Then,

$$f_z(\omega) = f_w(\omega) \{2 \sin(\omega/2)\}^{-2d}, \quad 0 \leq \omega \leq \pi \tag{2}$$

Where

$$f_W(\omega_j) = \frac{\sigma_a^2}{2\pi} \left| \frac{\theta_q(\exp(-i\omega_j))}{\phi(\exp(-i\omega_j))} \right|^2,$$

is the spectral density of a regular ARMA (p, q) model. Note that $f_Z(\omega) \rightarrow \infty$ as $\omega \rightarrow 0$.

b) Take logarithms on both sides of equation (2).

$$\begin{aligned} \ln\{f_Z(\omega_j)\} &= \ln \left| \frac{\theta_q(\exp(-i\omega_j))}{\phi(\exp(-i\omega_j))} \right|^2 + \ln f_W(\omega_j) \\ &= \ln f_W(0) + d \ln \left| \frac{\exp(i\omega_j)}{1 - \exp(i\omega_j)} \right|^2 + \ln \left(\frac{f_W(\omega_j)}{f_W(0)} \right) \end{aligned} \quad (3)$$

c) Add $\ln I_Z(\omega_j)$, the natural logarithm of periodogram $\{Z_t\}$ to Both sides of equation (3) above,

$$\ln\{I_Z(\omega_j)\} = \ln f_W(0) + d \ln \left| \frac{\exp(i\omega_j)}{1 - \exp(i\omega_j)} \right|^2 + \ln \left\{ \frac{f_W(\omega_j)}{f_W(0)} \right\} + \ln \left\{ \frac{I_Z(\omega_j)}{f_Z(\omega_j)} \right\} \quad (4)$$

d) Determine the periodogram from equation (4)

Geweke and Porter-Hudak (1983) obtained periodogram by

$$I_Z(\omega_j) = \frac{1}{2\pi} \left\{ \sum_{t=1}^{T-1} \cos(t\omega_j) \right\}^2, \quad \omega_j \in (-\pi, \pi). \quad (5)$$

Hurvich and Ray (1995) and Velasco(1999a) used modified periodogram function by

$$I_Z(\omega_j) = \frac{1}{2\pi \sum_{t=0}^{T-1} (\text{tap}(t))^2} \left| \sum_{t=0}^{T-1} \text{tap}(t) \exp(-i\omega_j t) \right|^2 \quad (6)$$

where the tapered data is obtained from the cosinebell function

$\text{tap}(t) = \frac{1}{2} \left[1 - \cos \left(\frac{2\pi(t+0.5)}{T} \right) \right]$. For Comparison, we use another tapers, there are Hanning

taper $\text{tap}(t) = \frac{1}{2} \left[1 - \cos \left(\frac{2\pi(t+1)}{T+1} \right) \right]$ and Hamming taper $\text{tap}(t) = 0.54 - 0.46 \cos \left(\frac{2\pi t}{T-1} \right)$.

e) Estimate the differencing parameter (d)

From equation (4), for ω_j near zero, i.e., for $j = 1, \dots, m \ll (T/2)$ such that $m/T \rightarrow 0$ as $T \rightarrow \infty$, we have $\ln \left(\frac{f_W(\omega_j)}{f_W(0)} \right) \approx 0$. Thus,

$$Y_j = \beta_0 + \beta_1 X_j + a_j, \quad j = 1, 2, \dots, m$$

$\hat{\beta}_1 = \hat{d}$ could be estimated by Ordinary Least Square (OLS). For computation, with

Euler equation we have $Y_j = \ln I_Z(\omega_j) X_j = \ln \left(\frac{1}{4 \sin^2(\omega_j/2)} \right)$.

After determining the value of d , we estimate Autoregressive (AR) and moving average (MA) parameters. Autoregressive (ϕ) and Moving average (θ) parameters are estimated by maximum likelihood method (Sowell, 1992).

4. SIMULATION STUDY

We have conducted simulation studies to obtain some information about the performance of the accuracy of spectral regression methods in estimating the degree of differencing parameter from ARFIMA model. In this simulation study we use three estimators five methods of spectral regression methods namely Cosine Bell, Hanning and Hamming tapers. The simulation is done for considering $T = 300, 600$ and 1000 serial data with 1000 replications. In this study, time series data are generated according to the particular specification Generate ARFIMA $(1,d,0)$ and ARFIMA $(0,d,1)$ models for simulated data, with $T = 600$ and $T = 1000$, respectively. Both types of data above have specification as follow $a_t \sim N(0,1)$, $(\phi, \theta) = 0.5$, $d = 0.6, 0.7$ and 0.8 .

For each series, we estimate the value of d through the three methods above and later we take the arithmetic average and standard deviation (SD) of these values.

After estimating ARFIMA parameters, we calculate Mean square error of forecasting with the value of out sample $h = 10$. MSE of forecasting could be seen at table 2.

Table 1 Mean and Standard deviation of parameter estimation d from ARFIMA model data.

T	ARFIMA Model Data		GPH Estimation Method With Cosine Taper					
			Cosine Bell		Hanning		Hamming	
			d	sd(d)	d	sd(d)	d	sd(d)
600	ARFIMA(1, d, 0)	d=0,6	0,624	0,195	0,620	0,200	0,617	0,195
		d=0,7	0,717	0,196	0,731	0,208	0,728	0,192
		d=0,8	0,840	0,198	0,833	0,193	0,830	0,189
	ARFIMA(0, d, 1)	d=0,6	0,581	0,201	0,570	0,205	0,585	0,192
		d=0,7	0,687	0,198	0,685	0,204	0,696	0,187
		d=0,8	0,791	0,203	0,790	0,201	0,787	0,191
1000	ARFIMA(1, d, 0)	d=0,6	0,611	0,172	0,602	0,173	0,6111	0,168
		d=0,7	0,712	0,178	0,705	0,170	0,716	0,197
		d=0,8	0,803	0,179	0,815	0,172	0,834	0,202
	ARFIMA(0, d, 1)	d=0,6	0,598	0,175	0,583	0,174	0,592	0,162
		d=0,7	0,695	0,172	0,677	0,176	0,687	0,187
		d=0,8	0,803	0,167	0,794	0,176	0,783	0,199

Table 2 MSE of Forecasting ($h = 10$)

T	ARFIMA Model Data		Cosine Taper		
			Cosine Bell	Hanning	Hamming
300	ARFIMA(1, d , 0)	$d=0,6$	1.316	1.317	1.311
		$d=0,7$	1.322	1.322	1.316
		$d=0,8$	1.342	1.342	1.336
	ARFIMA(0, d , 1)	$d=0,6$	1.254	1.254	1.251
		$d=0,7$	1.254	1.255	1.251
		$d=0,8$	1.232	1.233	1.229
600	ARFIMA(1, d , 0)	$d=0,6$	1.367	1.367	1.360
		$d=0,7$	1.323	1.323	1.317
		$d=0,8$	1.352	1.352	1.346
	ARFIMA(0, d , 1)	$d=0,6$	1.262	1.262	1.258
		$d=0,7$	1.219	1.219	1.215
		$d=0,8$	1.633	1.633	1.622

5. CONCLUSION

From simulation results, GPH method with Cosine Bell tapering shows a good performance in estimating the differencing parameter of ARFIMA(0, d , 1) data. From ARFIMA(1, d , 0) data, GPH method with Hanning tapering is the best of all methods. From forecasting result, Mean square error of GPH method with Hamming tapering has the least value of all data types.

6. ACKNOWLEDGEMENTS

I am grateful to Statistics Department and Faculty of Mathematics and Natural Science Padjadjaran University for financial support.

REFERENCES

- Bisaglia, L and Guegan, Dominique. (1998), "A Comparison of Techniques of estimation in long-memory processes", *Computational Statistics & Data Analysis*, Vol.27, p. 61-81.
- Fox, R and Taqqu, M.S. (1986). "Large-sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series", *The Annals of Statistics*, Vol.14, p. 517-532.
- Geweke J and Porter-Hudak,S. (1983), "The Estimation and Application of Long Memory Time Series Models", *Journal of Time series Analysis*, Vol. 4, p. 221-238.
- Granger, C. W. J. and Joyeux,R. (1980), "An Introduction to Long-Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis*, Vol. 1, p. 15-29.
- Hosking, J.R.M. (1981), "Fractional Differencing", *Biometrika*, Vol. 68, p. 165-176.
- Hurvich, C.M. and Deo,R.S. (1999), "An Introduction to Long-Memory Time Series Models and fractional Differencing", *Journal of Time series Analysis*, Vol. 20, p.331-341.
- Hurvich, C.M. and Ray, B.K. (1995), "Estimation of the Memory Parameter for Non stationary or Noninvertible Fractionally Integrated Processes", *Journal of Time series Analysis*, Vol. 16, p.17-42.
- Reisen, V.A. (1994), " Estimation of the Fractional Parameter for ARIMA(p,d,q) Model Using the Smoothed Periodogram", *Journal of Time Series Analysis*, Vol.15, p. 335-350.
- Reisen, V.A and Lopes,S.R.C. (1999), "Some Simulations and applications of forecasting long-memory time series models", *Journal of Statistical Planning and Inference*, Vol.80, p. 269-287.
- Robinson, P.M. (1995), "Log-Periodogram Regression of Time Series with Long Range Dependence", *Annals of Statistics*, Vol. 23, pp. 1048-1072.
- Sowell, F. (1992), "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models", *Journal of econometrics*, Vol.53, p.165 – 188.
- Velasco, C. (1999a), "Non-Stationary Log-Periodogram Regression", *Journal of Econometric*, Vol. 91, p. 325-371.
- Velasco, C. (1999b), "Gaussian Semiparametric Estimation of Non-Stationary Time Series", *Journal of Time Series Analysis*, Vol. 20, No.1, p. 87-127.