A NEW METHOD FOR ESTIMATING THE ORDER OF INTEGRATION OF FRACTIONALLY INTEGRATED PROCESSES USING BISPECTRA

May 3rd, 2005

Mehmet Dalkir
The University of Kansas,
Department of Economics
1300 Sunnyside ave.,
213 Summerfield Hall, Lawrence KS 66045
e-mail: mdalkir@ku.edu

Abstract

The method proposed in this chapter is making use of the bispectrum transformation to estimate the level of integration of a fractionally integrated time series. Bispectrum transformation transforms the series into a two dimensional frequency space, and thus has higher information content compared to the Geweke-Porter-Hudak method. The bispectrum method is an alternative to the recently proposed wavelet method that transforms the original series into time-frequency (or time-scale) space.

Keywords: Bispectrum, frequency domain, estimation, long memory.

1. Introduction

The first formal definition of a long memory process has been made by the British hydrologist Hurst (1951) on continuous time domain. Manderbolt et al. later defined the process in discrete time. Since then, long memory processes have been attracting great interest for financial and economic modeling (e.g., Baille, 1996). The long memory property in a time series could be defined loosely as persistence in the observed autocorrelations. The autocorrelations of a long memory process are more persistent than the autocorrelations of a typical stationary ARMA process, which gained much popularity among economists. The impulse response functions, or the coefficients of the Wold decomposition of a typical stationary ARMA process, together with their autocorrelations, display exponential decay. However, there is no reason to limit the decay to exponential rates. Other options like hyperbolic decay could also be considered. The integrated series concept partially filled this area, leaving still a big gap between series integrated of order zero and of order 1. The introduction of fractional integration aimed at filling this gap. Long memory processes describe financial time series like inflation rates, interest rate differentials, and volatility of asset prices quite well. The method conjectured here is an alternative to the relatively new methods including the estimation using wavelets. The wavelet transform could be defined in two dimensions: scale (frequency) and time. The bispectrum method, likewise, could be defined in a two dimensional frequency domain. So these two methods reveal information on the series in a two dimensional framework unlike the Geweke and Porter-Hudak (GPH) (1983). Therefore the bispectrum method proposed here is an alternative to the better known wavelet methods of estimating the fractional integration parameter.

2. Definition

McLeod and Hipel (1978) describe long memory with the sum of absolute autocorrelations (ρ_j) of a time series. According to their definition, a time series possess long memory property if the sum $\sum_{j=-n}^{n} |\rho_j|$ is divergent as $n \to \infty$. Equivalently, the spectral density $f(\omega)$ is divergent as $\omega \to 0$. A representation of long memory property that attracted particular interest is the fractionally integrated time series, which could be defined as

$$(1-L)^d y_t = u_t, \quad d \in \mathbb{R}, \tag{1}$$

where series $\{y_t\}$ is said to be integrated of order d, or is $I(d)$. When $\{u_t\}$ is white noise with variance σ^2, $\{y_t\}$ is said to be fractional white noise. The series above could be represented as an $AR(\infty)$ or $MA(\infty)$ process, where the AR coefficients will be

$$\pi_k = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)}, \tag{2}$$

or

$$\pi_k = \pi_{k-1}\frac{k-1-d}{k}, \tag{3}$$

and $\pi_o = 1$ where $\Gamma(\cdot)$ is the gamma function.
The MA representation is similar with a replacement of $-d$ for d in the AR representation. The autocorrelation functions take the form

$$\rho_k = \frac{\Gamma(k+d)\Gamma(1-d)}{\Gamma(k-d+1)\Gamma(d)}, \tag{4}$$

or

$$\rho_k = \frac{k-1+d}{k-d}\rho_{k-1}, \tag{5}$$

and $\rho_o = 1$.

The power spectrum is

$$f(\omega) = \frac{\sigma^2}{2\pi}(2\sin(\frac{\omega}{2}))^{-2d}. \tag{6}$$

Notice that $f(\omega) < \infty$ as $\omega \to 0$ iff $d \leq 0$, and since for small values of x, $\sin x \approx x$, $f(\omega) \approx \frac{\sigma^2}{2\pi}\omega^{-2d}$ for ω close to zero. Using Stirling's formula for the gamma function for large k, we can derive the asymptotic functions for the magnitudes discussed above. Stirling states that, for large k,

$$\frac{\Gamma(k+a)}{\Gamma(k+b)} \approx k^{a-b}. \tag{7}$$

So, when k is large,

$$\rho_k \approx \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1}, \tag{8}$$

and

$$\pi_k \approx \frac{k^{-d-1}}{\Gamma(-d)}, \tag{9}$$

which satisfy the hyperbolic (or geometric) decay condition with $-1 < d$ for π_k and with $d < \frac{1}{2}$ for ρ_k.

Granger and Joyeux (1980) and Hosking (1981) show that the series $\{y_t\}$ is also stationary for $d < \frac{1}{2}$. Odaki (1983) shows that $\{y_t\}$ is invertible if $-1 < d$. Thus $0 < d < \frac{1}{2}$ guarantees a stationary series with hyperbolically decaying autocorrelations, which satisfy the long memory condition.
In the literature, the focus is on the stationary side of $I(d)$ processes with $0 < d < \frac{1}{2}$. Over the region $\frac{1}{2} \leq d < 1$, there is still a decay in the impulse response functions and the power spectrum as frequencies go low. However, the autocorrelations get higher as k grows, which puts the process into the nonstationary processes group [1] .

3. Estimation

The estimation procedure for the long memory series in this study is concentrated on the method proposed by GPH (1983). In their paper, they exploit the spectral density functional form of a stationary long memory series to derive an estimate of the long memory parameter d.
The spectral density function for the series $\{X(t)\}$ where $(1-L)^d X_t = u_t$ and $\{u_t\}$ is a stationary linear process with spectral density function $f_u(\omega)$ is

$$f(\omega) = \frac{\sigma^2}{2\pi}(4\sin^2(\frac{\omega}{2}))^{-d} f_u(\omega). \tag{10}$$

[1] For a better explanation of stationarity of $I(d)$ processes, see Hamilton (1994).

Manipulating the above equation and taking logarithms of both sides of the equation, they obtain

$$\ln I(\omega_{j,T}) = \ln \frac{\sigma^2}{2\pi} f_u(0) - d \ln 4 \sin^2(\frac{\omega_{j,T}}{2}) + \ln \frac{f_u(\omega_{j,T})}{f_u(0)} + \ln \frac{I(\omega_{j,T})}{f(\omega_{j,T})}, \quad (11)$$

which is in the form of a linear regression model. GPH argue that if frequencies close to zero are considered, the third term on the right hand side of the equation will be negligible, the second term will be the independent variable. The last term (less its mean) is the disturbance, and the first term of the right hand side plus the mean of the last term constitutes the constant of the regression. GPH prove in their Theorem 2 (Geweke and Porter-Hudak, 1980) that when $d < 0$ there exists a function of the sample size $g(T) < T$, such that if the smaller $n = g(T)$ frequencies are used for the estimation of the regression model above, the OLS estimator of the slope will consistently estimate $-d$. They generalize this result empirically for $0 < d < \frac{1}{2}$.

4. A proposed estimation process using bispectra

It is well known that (see e.g., Sakaguchi and Sakai, 1989) all polyspectra of order higher than 2 of a Gaussian process are zero. Furthermore, if the process is linear, then,

$$\frac{f_3^2(\omega_1, \omega_2)}{f(\omega_1)\ f(\omega_2)\ f(\omega_1 + \omega_2)} = constant, \quad \forall \omega_1, \omega_2, \quad (12)$$

where $f_3(\omega_1, \omega_2)$ is the bispectrum of the process. Hinich (1982) uses this property to design a test to identify Gaussianity and linearity of a time series from its bispectrum.
Consider the time series $\{X(t)\}$ above, where $\{u_t\}$ is non-Gaussian. The bispectrum of $\{X(t)\}$ could be written as

$$f_3(\omega_1, \omega_2) = h(\omega_1) h(\omega_2) h(-\omega_1 - \omega_2) Q(\omega_1, \omega_2), \quad \forall \omega_1, \omega_2, \quad (13)$$

where $h(\omega)$ is the transfer function for any linear filter, and Q is the bispectrum of process $\{u_t\}$. Since the lag polynomial $(1 - L)^d$ of an $I(d)$ series is a linear filter with transfer function

$$h(\omega) = |2 \sin(\omega/2)|^{-d}, \quad (14)$$

the above property could be used to estimate d, in a similar fashion with GPH. Manipulating the equation and taking logarithms of both sides, we obtain

$$\begin{aligned}\ln\ I_3(\omega_1,\omega_2) &= \ln\frac{I_3(\omega_1,\omega_2)}{f_3(\omega_1,\omega_2)}+\ln\frac{\sqrt{f_u(\omega_1)f_u(\omega_2)f_u(-\omega_1-\omega_2)}}{f_u^3(0)} \\ &+\ln\ f_u^3(0)-d\ln|2\sin\frac{\omega_1}{2}|-d\ln|2\sin\frac{\omega_2}{2}|-d\ln|2\sin\frac{\omega_1+\omega_2}{2}|,\end{aligned} \tag{15}$$

where $I_3(\omega_1,\omega_2)$ is a consistent estimator of bispectrum of $X(t)$.

The argument of GPH holds here: the first term on the right hand side of the equality is the disturbance, the second term is negligible, the third term is constant, and the last three terms represent independent variables of the regression equation:

$$y_{ij}=\phi_o+\phi_1 h_{i0}+\phi_2 h_{0j}+\phi_3 h_{ij}+v_{ij}, \tag{16}$$

where $1\leq i\leq j\leq T$, and $h_{ij}=\ln|2\sin(\frac{i+j}{T}\pi)|$. The disturbance is represented by the term v_{ij}.
One advantage of the bispectrum representation over the power spectrum representation is that now the conclusion of the estimation process is testable. If the null hypothesis

$$H_o:\phi_1=\phi_2=\phi_3 \tag{17}$$

is rejected, then the estimation procedure is inconclusive. (Note that $\phi_1=\phi_2=\phi_3=-d$). A model with a smaller number of parameters could be

$$y_{ij}=\beta_o+\beta_1(h_{i0}+h_{0j})+\beta_2 h_{ij}+w_{ij}. \tag{18}$$

Again, the test for conclusiveness would be to test the null

$$H_o:\beta_1=\beta_2. \tag{19}$$

4.1. Asymptotic consistency

Brillinger and Rosenblatt (1967) show that estimates of any k^{th} order spectra are asymptotically normal. However, since all the variables in the model are in logarithms, and other inferences from the independent variables are uncertain, it is hard to say that the disturbance terms are asymptotically normal. However, as explained in Greene (2000, Chapter 11) maximum likelihood estimation of the coefficients will yield consistent estimators once Gaussian disturbances are assumed. (Hence the estimation method is quasi-maximum likelihood.) The consistent covariance matrix for the parameters could be estimated through the "sandwich" estimator (White, 1982). The result of any test under consistency will thus be reliable for a large number of observations. For the test of conclusiveness, the maximum likelihood ratio test could be used to compare the likelihood function values of one of the regression equations above with that of

$$y_k = a_o + a_1(h_{i0} + h_{0j} + h_{ij}) + \xi_{ij}. \tag{20}$$

If the estimation procedure is found to be conclusive, the quasi-maximum likelihood estimator $-a_1^{QML}$ could be taken as a consistent estimator of d.

5. Estimation with wavelets

Wavelets shed new light on many problems in time series analysis, including estimation of fractional differencing parameters. Wavelet transformations have the unique property of transferring a series from time domain to a two dimensional time-scale (or time-frequency) domain. It is this property that combines the estimation in time domain and frequency domain on common grounds. The estimation of the fractional integration parameter is accomplished through a plot of the wavelet variances (Percival and Walden, 2000, section 9.5) versus the scale of estimation. The procedure is then to find the part of the plot that visually appears to be linear, and estimation of the slope of that line. The problem of which observations to include prevails here in terms of wavelet variances in different scales.

6. Conclusion

Almost for half a century, the fractional integration representation of long memory processes has attracted attention from a wide variety of disciplines and researchers. The convenience of the fractionally integrated model representation is the one and only parameter to be estimated. Still the task is far from being trivial. Many estimation procedures were later shown to be systematically over- or underestimating the model parameter. One advantage of the bispectrum method of estimation is that the conclusiveness of the procedure can be tested. If the test result implies a convenient estimator, then one further step is taken to identify the estimator. If the test fails, however, another method of estimation in the literature could be employed.

The method conjectured here is an alternative to the relatively new methods including the estimation using wavelets. The wavelet transform could be defined in two dimensions: scale (frequency) and time (see e,g., Gençay, Selçuk, and Whitcher, 2002). The bispectrum method, likewise, could be defined in a two dimensional frequency domain. So the two methods reveal information on the series in a two dimensional framework unlike most other estimation methods. There are a few remaining issues to be considered while working with the procedure:

- The disturbance of the underlying process should be non-Gaussian. If it is Gaussian, the procedure will not work since the bispectrum collapses to zero for all frequencies in this case.
- The model used for estimation is not a classical regression model. The deviation of the estimators from consistency should be found using Monte Carlo simulation techniques. Likewise, the estimated "sandwich" covariance matrix may not be reliable, especially in the case where the disturbances are fat-tailed.
- The question "How close should the frequencies to be used be close to zero?" from the procedure of GPH prevails in this method of estimation. A bispectrum version of their function $g(T)$ should be derived, or the efficiency of various filters should be tested on simulated data.
- The effectiveness of the procedure for the nonstationary case, (when d is greater than $\frac{1}{2}$) is crucial since most methods in literature leave nonstationary series out of their scope.
  Further work is warranted to address these issues.

REFERENCES

Baille R. (1996) "Long memory processes and fractional integration in econometrics" *Journal of Econometrics*, **73**, 5-59.

Blanchard, O., and Simon, J. (2001), "The Long and Large Decline in U.S. Output Volatility," *Brookings Papers on Economic Activity* vol.**1**, pp. 135-164.

Brillinger D.R., and Rosenblatt M. (1967) "Asymptotic theory of estimates of the k-th order spectra" in: Harris, B. (Ed.) *Spectral Analysis of Time Series*, Wiley.

Brillinger, D.R. (1975) *Time Series: Data Analysis and Theory* Holt, Rinehart and Winston, Inc.

Geweke J., and Porter-Hudak S. (1983) "The estimation and application of long memory time series models", *Journal of Time Series Analysis*, **4**, 4.

Geweke, J., Meese, R., and Dent, W., (1983) "Comparing alternative tests of causality in temporal systems," *Journal of Econometrics,* **21**, 161-194.

Granger C., and Joyeux R. (1980) "An introduction to long-memory time series models and fractional differencing," *Journal of Time Series Analysis*, **1**, 15-29.

Hamilton J. (1994) *Time Series Analysis*, Princeton University Press.

Hinich M.J. (1982) "Testing for gaussianity and linearity of a time series," *Journal of Time Series Analysis*, **3**, 169-76.

Hosking J.R.M. (1981) "Fractional differencing," *Biometrika*, **68**, 1, 165-76.

Koopmans, L.H. (1974) *The Spectral Analysis of Time Series* Monograph and Textbook Series in Probability and Mathematical Statistics, Academic Press: London.

McLeod A.I., and Hipel K.W. (1978) "Simulation procedures for Box-Jenkins models," *Water Resources Research*, **14**, 5, 969-75.

Odaki M., and Hatanaka M. (1983) "Policy analyses with and without a priori conditions," *Economic Studies Quarterly*, **34**, 3, 193-210.

Ramsey, J.B., and Lampart, C. (1998) "Decomposition of Economic Relationships by Timescale Using Wavelets: Money and Income" *Macroeconomic Dynamics* **2**, 49-71.

Rudebusch, G. D. (2003), "The uncertain unit root in real GNP," *American Economic Review,* 83(1), 265-72.

Sakaguchi F. and Sakai H. (1989) “A Composite linear model generating a stationary stochastic process with given bispectrum,” Workshop on Higher Order Spectral Analysis, Vail, Colorado, June 28-30.

White, H. (1982) “Maximum Likelihood Estimation of Misspecified Models,” *Econometrica*, **50**,1.