A linearity test for a class of processes with long-range dependence

Felipe Miguel Aparicio Acosta

Signal Processing Laboratory, Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland
Phone: +41.21.6934714 Fax: +41.21.6937600 E-mail: aparicio@ltssun2.epfl.ch

Abstract Long-Range Dependence (LRD) *arises in communications (1/f noises) [1], and in many other fields, such as hydrology [2], astronomy [3], meteorology [4], finance [5] and the economic sciences [6]. In many problems, we are given a time series from the output of a dynamical system whose behaviour we would like to characterize. For example, one may wish to know whether a system behaves linearly or not, under the assumption that its input is linear. However, when the output signal has LRD in the mean, it may be difficult to detect the nonlinearity. In this paper, we attempt to explain one of the reasons why this is so, and propose a linearity test to overcome this problem on a class of nonlinear time series with LRD. Its power is compared to that of some classical linearity tests.*

Résumé *Les signaux à* longue mémoire *apparaisent dans beaucoup de situations. On les retrouve aussi bien en communications (bruits 1/f) [1], qu'en sciences économiques [6], finances [5], astronomie [3], méteorologie [4] ou en hydrologie [2]. Dans certains problèmes, on s'interesse a savoir si un système dynamique se comporte linéairement ou non. Pour ce faire, on peut appliquer un test de linéarité à la série chronologique obtenue à partir de sa réponse. Il se trouve que lorsqu'il y a de longue dépendence dans la série, les tests de linéarité paramétriques deviennent très sensibles au choix de la fonction de test. Dans ce papier, on se propose d'expliquer ce problème et d'en proposer une solution basée sur un choix approprié de la fonction de test, pour une classe des séries a longue mémoire.*

1 Introduction

A process, x_t, is often said to be LRD *persistent* (in the mean), if it has non-summable autocorrelations, $\rho_x(\tau)$, verifying $\rho_x(\tau) \approx \alpha\tau^{-\beta}$, for large $|\tau|$ $(0 < \beta)$, and a locally unbounded spectral density, $S_x(\lambda)$, behaving like $|\lambda|^{\beta-1}$ at least one frequency location. This is in constrast with processes having *Short-Range Dependence* (SRD) in the mean, for which the spectral density is bounded everywhere and the autocorrelation function decays exponentially as the lag increases. Many dynamical systems act on LRD signals, while their outputs could be LRD or not. In applications such as system characterization or classification, an important issue may be to determine whether one or several systems behave linearly or not, or whether some behave more nonlinearly than others. Linearity testing is used to estimate the incidence of nonlinearity in a given time series, and may also help to infer the type of nonlinearity in the data by using a battery of *test functions*.

An important family of tests are those tailored against specific nonlinear alternatives. These tests postulate a model that encompasses both the null hypothesis of linearity and the alternative of nonlinearity. That is, a model such as

$$y_t = \theta_1{}' Y_{t-1,1} + g_{\theta_2}(Y_{t-1,2}) + \xi_t \tag{1}$$

where $\theta = (\theta_1, \theta_2)'$ is the parameter vector for the model, $Y_{t-1,1}$ and $Y_{t-1,2}$ are vectors of lagged values of the process (that may have different dimensions), and ξ_t is assumed to be a zero-mean Gaussian *i.i.d.* sequence. Under the null, both θ_2 and $g_{\theta_2}(.)$ are equal to zero. Therefore testing linearity amounts at testing for $\theta_2 = 0$ in this nested model.
The nonlinear function $g_{\theta_2}(.)$ is intended to describe the form of nonlinearity in the data, and it is assumed to be at least twice continuously differentiable on the parameter vector.
Most parametric linearity tests can be put into this framework, as shown in [7]. Take for example, those inspired by Tukey's one-degree of freedom test for non-additivity [8], such as Keenan's [9] and Tsay's [10]. Both test for the significance of the interaction or nonlinear terms in a second-order Volterra expansion of y_t,

$$y_t \approx \mu + \sum_{i=1}^{\infty} b_i \epsilon_{t-i} + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} b_{i,j}\epsilon_{t-i}\epsilon_{t-j} + \epsilon_t \qquad (2)$$

They rely on the assumption that nonlinearity in y_t would imply that $b_{i,j}$ are significantly different from 0, for at least some i, j. This nonlinearity would be reflected in the diagnostics of a linear model fitted to y_t, because the residuals should then be correlated with second-order cross-terms, $y_{t-i}y_{t-j}$.

2 Some problems of linearity testing under LRD

Nonlinearity can be masked by persistence in different ways [4]. One possibility is due to an *imbalance* of both sides in the nonlinear regression equation that usually forms the second stage of parametric linearity tests. This entails low-power for the test. For example, imagine the following model

$$y_t = ay_{t-1} + by_{t-1}\xi_{t-1} + \xi_t \qquad (3)$$

where $a \approx 1$ and ϵ_t is an *i.i.d.* sequence. This corresponds to a bilinear near-unit root time series. These series are LRD in the mean, in the sense that the conditional mean $E(y_{t+m}|y_t)$ never goes to zero as the forecasting horizon m goes to infinity. In fact, if we wish to stick to the definition of LRD given above, it is possible to show that they have an hyperbolic spectral behaviour in f^2 (of the form $1/f^2$) near the origin, and that their linearly decaying sample autocorrelation function (ACF) could be regarded as a degenerate hyperbola. Misspecification of the disturbances in the right-hand side of (3) would lead to a linear AR(1) model as the best fitting model. The resulting residuals would be SRD in the mean and cause an imbalance problem when regressed on a nonlinear function, $g(.)$, of y_{t-1}, that is itself LRD; for example, $g(y_{t-1}) = y_{t-1}^2$.

3 A linearity test for a class of processes with LRD

The imbalance problem appears when the linear residuals, $\epsilon_t = y_t - \theta_1' Y_{t-1,1}$, are SRD in the mean while the regressor generated by the testing functions are not. For example, if y_t is linear LRD in the mean so will be y_t^2 (this can be easily checked in the frequency domain, since the spectrum of y_t^2 results from convolving that of y_t with itself) (figure 1). Fortunately, there are nonlinear transformations of y_t for which LRD is absent or, at least, reduced to some extent. Take for instant such bounded transformations based on the exponential function as $exp(-\frac{|y_t|}{\sigma})$, $y_t exp(-\frac{y_t^2}{\sigma^2})$, or on trigonometric functions, such as $sin(y_t)$, or $cos(y_t)$ (see figure 3). Hallman [11] showed that the exponential transformation of a LRD-in-the-mean process such as a random walk, may no longer be LRD in the mean. In fact, for the random walk case, the transformed series has the same autocorrelation structure as a stationary AR(1) process, even though (strictly speaking) it is still LRD in a higher-order moment [12]. On the other hand, trigonometric transformations of a LRD-in-the-mean process, such as $y_t = sin(x_t)$, where x_t is LRD in the mean, have been shown to lead to stationary behaviour [13, 12]. In particular, if $y_t = sin(x_t)$, with x_t denoting a random walk, then y_t is a stationary AR(1) process. Therefore, in order to restore the balance in the regression equation of ϵ_t on $\psi_{\mathbf{a}}(Y_{t-1})$, we need to construct a sufficiently mixing process z_t, from an appropriate test function $z_t = \psi_{\mathbf{a}}(Y_t)$, which will restore the balance of the nonlinear regression equations in a parametric test. A situation where such an strategy has proven succesful is in testing against *near-unit root processes* with nonlinearity in the mean or with conditional heteroskedasticity. These processes may arise in the output of some nonlinear sensor devices measuring a LRD (in the mean) flow [4]. A most simple testing function that is appropriate for these nonlinear processes, and which we have used in our test, is $g_\theta(Y_{t-1}) = y_{t-1}\exp[-(Y_{t-1} - C)'\Sigma(Y_{t-1} - C)]$, where Σ is a positive definite matrix, and $\theta = (C, \Sigma)$. The spectral contents of ϵ_t and $g_\theta(Y_{t-1})$ can be matched to a good degree of accuracy by tuning the parameters in Σ. This smoothing matrix controls the amount of nonlinearity and creates the necessary mixing that destroys LRD (see figure 4).

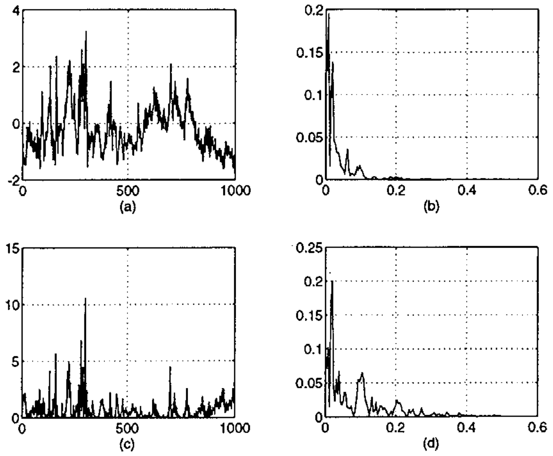

Figure 1: Preservation of LRD in the mean after a square trasformation: (a) real-world series with LRD in the mean, (b) its spectral density estimate showing the pole at the origin, (c) the square of the previous series, and (d) the spectral density estimate of (c), showing still the singularity near the origin.

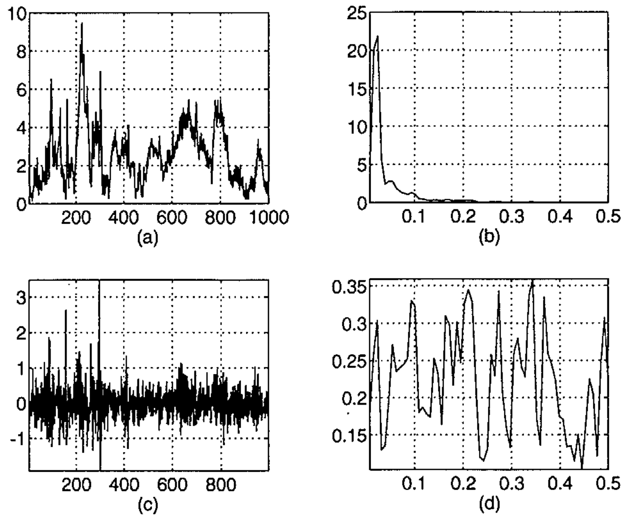

Figure 2: A real-world example that shows the imbalance problem in linearity testing. The spectral features of the series and its linear residuals are unmatched near the origin: (a) sample of wind speed data from a nonlinear sensor, (b) spectral density estimate of (a), (c) linear residuals from an optimal (using AIC) AR model fitted to (a), and (d) spectral density estimate of these residuals.

4 Results

In the experiments, we considered the model $y_t = a(y_{t-1})y_{t-1} + by_{t-1}\epsilon_t$. The power of our test ($T0$) was estimated using the asymptotic critical values at the 5% level, and compared to the power of both Keenan's ($T1$) and Tsay's ($T2$) tests, on 500 independent replications of near-unit root time series with and without conditional heteroskedastic dependencies. There were 500 observations from each time series, which were generated according to this model, with $a(y_{t-1})$ constant (model A), and with $a(y_{t-1}) = 0.9 + 0.1exp(-\sigma^2 y_{t-1}^2)$ (model B) (in both cases, at the border of the nonstationary region), and for different values of the smoothing parameter σ. As shown in the table below, our test outperformed the two others in all the cases.

MODEL versus *TEST*	*T0*	*T1*	*T2*
A $(b = 1.0)$	0.6	0.2	0.29
B $(b = 0.0, \sigma^2 = 1.0)$	0.15	0.03	0.04
B $(b = 1.0, \sigma^2 = 1.0)$	0.8	0.42	0.55
B $(b = 0.0, \sigma^2 = 10.0)$	0.5	0.04	0.02
B $(b = 1.0, \sigma^2 = 10.0)$	1.0	0.54	0.59
B $(b = 0.0, \sigma^2 = 100.0)$	0.55	0.02	0.08
B $(b = 1.0, \sigma^2 = 100.0)$	1.0	0.68	0.86

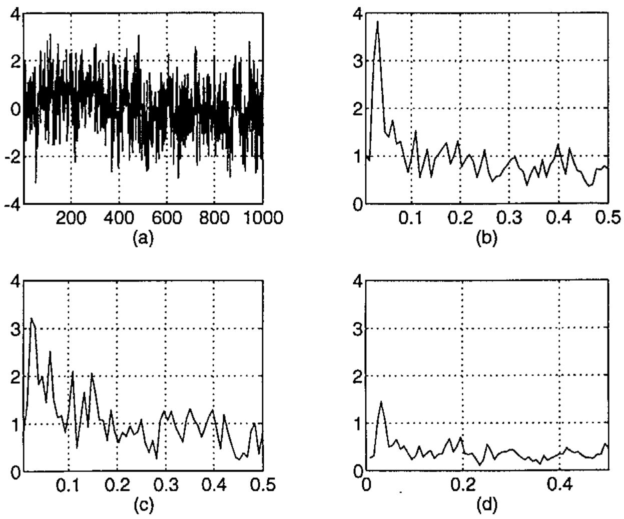

Figure 3: Periodogram estimates for a stationary fractionally differenced white-noise series, x_t, with long-memory parameter $d = 0.2$ (a)-(b), its exponential transformation $exp(x_t)$ (c), and its sinus transformation $sin(x_t)$ (d). The singularity of the spectral density at the origin is attenuated, if not removed, by these transformations.

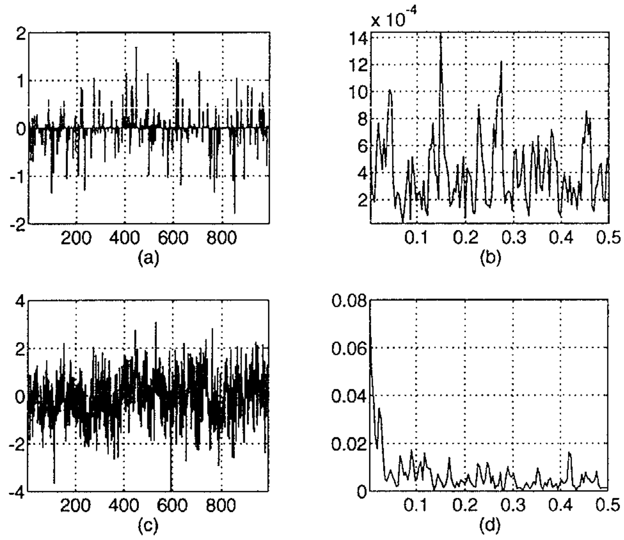

Figure 4: Periodogram estimates for the transformation $y_t = x_t\,exp(-\sigma^2 x_t^2)$, where x_t is a fractionally ARIMA$(0, d, 0)$ with long-memory parameter $d = 0.3$: (a) and (b) corresponds to the transformed series and its periodogram (respectively) for σ equal to 2 times the standard deviation of x_t, whereas (c) and (d) corresponds to a value of the smoothing parameter, σ, 10 times smaller. Notice how LRD in the mean re-appears progressively in the transformed series as the value of the smoothing parameter is decreased, that is, as the series becomes "more linear".

References

[1] A. Vand Der Ziel. Unified presentation of $1/f$ noise in electronic devices: fundamental $1/f$ noise sources. *Proceedings of the IEEE*, 76(3):223–258, March 1988.

[2] J.R.M. Hosking. Modeling persistence in hydrological time series using fractional differencing. *Water resources research*, 20(12):1898–1908, 1984.

[3] E.D. Feigelson and G.J. Babu. *Aperiodic time series in astronomy.* Statistical challenges in modern astronomy. Springer-Verlag, N.Y., 1992.

[4] F.M. Aparicio Acosta. *Nonlinear modelling and analysis under long-range dependence with an application to positive time series.* PhD thesis, Swiss Federal Institute of Technology (EPFL) -Signal Processing Laboratory-, Lausanne, Switzerland, 1995.

[5] E.E. Peters. *Fractal market analysis.* John Wiley and Sons, N.Y., 1994.

[6] C.W.J. Granger. The typical spectral shape of an economic variable. *Econometrica*, 34:150–161, 1966.

[7] R. Luukkonen, P. Saikkonen, and T. Terasvirta. Testing linearity in univariate time series models. *Scandinavian Journal of Statistics*, 15:161–175, 1988.

[8] J.W. Tukey. One degree of freedom for non-additivity. *Biometrics*, 5:232–242, 1949.

[9] D.M. Keenan. A Tuckey nonadditivity-type test for time series nonlinearity. *Biometrika*, 72(1):39–44, 1985.

[10] R.S. Tsay. Nonlinearity tests for time series. *Biometrika*, 73:461–466, 1986.

[11] J.J. Hallman. *Nonlinear integrated series, cointegration, and an application.* PhD thesis, Department of Economics of the Univ. of California at San Diego, La Jolla, USA, 1990.

[12] L. Ermini and C.W.J. Granger. Generalizations on the algebra of I(1) processes. *Journal of econometrics*, 58:369–384, 1993.

[13] C.W.J. Granger and J.J. Hallman. The algebra of I(1) processes. In *Finance and economic discussion series*, volume 45. Board of Governors of the Federal Reserve System, Washington D.C., 1988.