

Adaptive Estimation of Frational Exponential Models ¹

Stima Adattiva per Modelli Esponenziali a Differenze Frazionarie

Marcella Corduas

Dipartimento di Scienze Statistiche, Università di Napoli Federico II

e-mail:corduas@unina.it

Riassunto: In molti campi di applicazione sono rilevate serie storiche che presentano componenti a memoria lunga. In questo lavoro viene esaminata la classe dei modelli esponenziali a differenza frazionaria (FEXP) e viene proposto un criterio per la stima adattiva che minimizza la somma dei quadrati degli errori di previsione ℓ -passi avanti. Tale procedura di stima può realizzare, laddove il modello generatore della serie non è noto, un miglioramento delle previsioni prodotte.

Keywords: FEXP models, Adaptive estimation, Multi-step forecasting.

1. Introduction

We consider a Gaussian long-range dependent process X_t with spectral density function:

$$f(\omega) = |1 - e^{-i\omega}|^{-2d} f^*(\omega), \quad \omega \in (-\pi, \pi] \quad (1)$$

where $0 \leq d < 0.5$ is the parameter that controls the long-term persistence and $f^*(\omega)$ is the spectral density function of a stationary and invertible short-memory process.

Many empirical time series arising, for instance, in economics, hydrology, computer science, present a spectral density function with this form. In the literature of the past decade there has been a rapidly increasing interest in studying estimating procedures for the long memory parameter. Contributions have been developed in the parametric context deriving the properties of maximum likelihood estimator (MLE) both in frequency and time domain, and in the semiparametric context considering mainly methods based on log-periodogram regression (see Beran, 1994, for a review).

In this paper we examine Fractional Exponential models (FEXP) initially discussed by Janacek (1982) and Beran (1993) as an extension of the models introduced by Bloomfield (1973) for short-memory time series. Recently, Moulines and Soulier (2000), Hurvich (2001), Hurvich and Brodsky (2001) have reconsidered this class of models investigating the statistical properties of the fractional differencing estimator defined by least squares regression in spectral domain.

In the following section, moving from a parametric approach we propose an adaptive estimation procedure for FEXP models based on minimizing the multi-step ahead forecast error sum of squares. The use of a multi-step estimation criterion can improve forecasting accuracy, as discussed by Tiao and Xu (1983), Tiao and Tsay (1994) when the true model is unknown. In fact, we recall that a linear model for Gaussian time series is usually estimated by minimizing the sum of squares of one-step ahead forecast errors and

¹This research was supported by M.I.U.R. and C.N.R. and in part was carried out at the Centro per la Formazione in Economia e Politica dello Sviluppo Rurale, Portici.

the estimated parameters are used to determine future values for all time period of interest. This method provides asymptotically ML estimators, but it yields optimal efficiency for predictive application only if the model is correct. This situation is in practice rarely verified since, of course, all fitted models are false. At most, they are approximation of the unknown true data generating process. Thus it is worth investigating an estimation criterion which takes into account the prediction objective and adjusts the parameter estimates as a function of the forecast horizon.

2. An Adaptive Estimation Criterion

A FEXP model is formulated explicitly in terms of the spectral density function (1) where $f^*(\omega)$ is given by the exponential model:

$$f^*(\omega) = \frac{\sigma^2}{2\pi} \exp(2 \sum_{j=1}^k \delta_j \cos(j\omega)), \quad \omega \in (-\pi, \pi] \quad (2)$$

Since X_t is a stationary process it has a one-sided moving average representation: $X_t = (1 - B)^{-d} \psi^*(B) a_t = \psi(B) a_t$, where using standard notation: $a_t \sim WN(0, \sigma^2)$ is a White Noise process with zero mean and variance σ^2 , $\psi(B) = 1 + \sum_{j=0}^{\infty} \psi_j B^j$ with $\sum \psi_j^2 < \infty$ and B is the backward operator. Similarly, $\psi^*(B)$ is defined. We notice that the ψ_j^* are related to the δ_j coefficients by the following relation: $\exp(\sum \delta_j z^j) = (1 + \sum \psi_j^* z^j)$, being z a complex variable (Bloomfield, 1973).

Thus, the multi-step ahead prediction of $X_{t+\ell}$ based on the information up to time t is: $X_t(\ell) = \sum_{j=0}^{\infty} \psi_{t+j} a_{t-j}$ and, letting $\psi_{\ell-1}(B) = \sum_{j=0}^{\ell-1} \psi_j B^j$, the related ℓ -step ahead forecast error is: $\eta_t(\ell) = \sum_{j=0}^{\ell-1} \psi_j a_{t+\ell-j} = \psi_{\ell-1}(B) a_{t+\ell}$.

Referring to the information up to time $t - \ell$, the ℓ -step ahead forecast error can be easily expressed as: $\eta_{t-\ell}(\ell) = \psi_{\ell-1}(B) a_t = \psi_{\ell-1}(B) \psi^{-1}(B) X_t$. Of course, when $\ell = 1$, the one-step ahead forecast error is simply $\eta_{t-1}(1) = a_t$.

In this case, given an observed Gaussian time series $\mathbf{x} = \{x_1, x_2, \dots, x_n\}'$, when n is fairly large, the logarithm of the likelihood of the data can be approximated by a least-squares criterion:

$$Q_n(\mathbf{x}, \delta_1, \dots, \delta_k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{t=1}^n x_t e^{-i\omega t} \right|^2 [g(\omega, \delta_1, \dots, \delta_k)]^{-1} d\omega \quad (3)$$

where $g(\omega) = f(\omega) 2\pi / \sigma^2$. The estimates $(\hat{\delta}_1, \dots, \hat{\delta}_k)$ are obtained by minimizing (3) and setting $\hat{\sigma}^2 = Q_n(\mathbf{x}, \hat{\delta}_1, \dots, \hat{\delta}_k) / n$. When the Gaussian assumption is relaxed the above method, although no longer an approximate ML method, provides still estimates that possess most of the expected asymptotic properties (see Walker, 1964).

Extending this argument, we define the ℓ -step ahead forecast error sum of squares: $Q_n(\mathbf{x}, \delta_1(\ell), \dots, \delta_k(\ell)) = \sum_t \eta_{t-\ell}^2(\ell)$, which, neglecting end-effects, can be formulated in the frequency domain (Priestley, 1981, pp. 422-423; Haywood and Tunnicliffe Wilson, 1997):

$$\begin{aligned} \sum_t \eta_{t-\ell}^2(\ell) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_t \eta_{t-\ell}(\ell) e^{-i\omega t} \right|^2 d\omega = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\psi_{\ell-1}(e^{-i\omega})|^2 \left| \sum_{t=1}^n x_t e^{-i\omega t} \right|^2 [g(\omega, \delta_1, \dots, \delta_k)]^{-1} d\omega. \end{aligned} \quad (4)$$

Notice that the criterion (3) is modified by the gain function: $G_\ell(\omega) = |\psi_{\ell-1}(e^{-i\omega})|^2$ which is simply equal to one when $\ell = 1$. In real applications, the integral (4) is approxi-

mated as:

$$\sum_{j=1}^{[(n-1)/2]} G_\ell(\omega_j) I_n(\omega_j)/g(\omega_j), \quad (5)$$

where $I_n(\omega_j) = (2/n) |\sum_{t=1}^n x_t e^{-i\omega t}|^2$ is the time series periodogram and $\omega_j = 2\pi j/n$, $j = 1, \dots, [(n-1)/2]$ are the Fourier frequencies. The contribution of $I_n(\omega_j)/g(\omega_j)$ to the estimation criterion at each frequency is then corrected by a weight provided by $G(\omega_j)$, as it will be esemplified in the next section for a model fitted to a simulated time series.

3. A Monte Carlo Experiment

To evaluate the performance of the proposed estimation criterion we conducted a Monte Carlo experiment assuming that the 'true' generating process is ARFIMA. We generated 100 independent time series with $n = 800$ from a Gaussian ARFIMA(1,d,0) model with $d = 0.35$ and the Autoregressive parameter $\phi = 0.5$. We considered a FEXP model with $k=1$ as an adequate approximation and estimated the parameters (d, δ_1) using as criteria: the sum of squares of errors corresponding to one-step ahead and the ℓ -step ahead forecasts, respectively. Then, each time series was predicted for 20 time points.

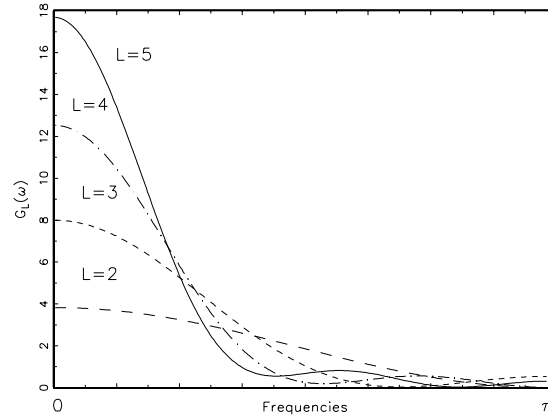


Figure 1. Gain functions for $\ell=2,3,4,5$ (lower curve $\ell = 2$, upper curve $\ell = 5$)

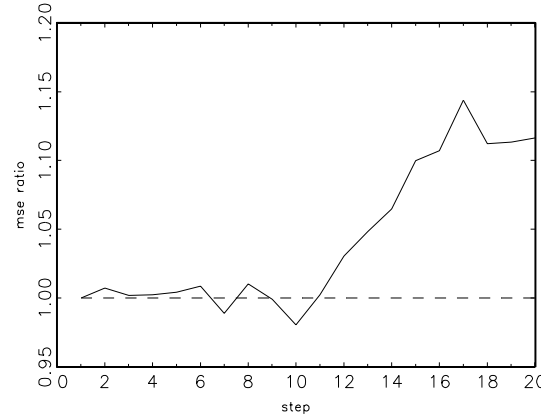


Figure 2. Ratio of mean square errors of post-sample forecasts: 1-step vs. ℓ -step

In Figure 1 we report the plot of $G_\ell(\omega)$ obtained for selected values of the step ℓ when the FEXP model is fitted to one of the simulated time series. For $\ell = 1$, $G_1(\omega) = 1$ whereas for $\ell = 2$ we obtain the lower curve in the plot. As ℓ increases the gain function assumes higher values about zero. The effect of the gain function on the estimation criterion seems reasonable since in order to adjust the criterion to farther time period forecasts more weight is given to lower frequencies that is to long term components.

In Figure 2 we report the ratio of the mean square errors of post-sample forecast from non adaptive estimation with respect to adaptive forecasting. It shows how, after the initial steps, the multi-step criterion ensures some improvements (up to 15%) in predicting accuracy as the forecast horizon increases.

The results confirm that for long memory time series adaptive estimation may improve the forecast ability of an uncorrect model. Further studies are in progress in order to assess other statistical properties of the proposed estimators.

References

- Beran, J. (1993) Fitting long-memory models by generalized linear regression, *Biometrika*, 80, 817-822.
- Beran J. (1994) *Statistics for Long-memory Processes*, Chapman & Hall, New York
- Bloomfield, P. (1973) An Exponential Model for spectrum of a scalar time series, *Biometrika*, 60, 217-226.
- Haywood J., Tunnicliffe Wilson G. (1997) Fitting time series models by minimizing multi-step ahead errors: a frequency domain approach, *Journal of the Royal Statistical Society, B*, 237-254.
- Hurvich C.M. (2001) Model selection for broadband semiparametric estimation of long-memory in time series, *Journal of Time Series Analysis*, 22, 679-709.
- Hurvich C.M., Brodsky, J. (2001) Broadband semiparametric estimation of long-memory time series using Fractional Exponential Model, *Journal of Time Series Analysis*, 22, 221-249.
- Janacek G.J. (1982) Determining the degree of differencing for time series via the log spectrum, *Journal of Time Series Analysis*, 3, 177-183.
- Moulines E., Soulier, P. (2000) Data driven order selection for projection estimator of the spectral density of time series with long range dependence, *Journal of Time Series Analysis*, 21, 193-218.
- Priestley, M.B. (1981) *Spectral Analysis and Time Series*, Academic Press, London.
- Tiao, G. C., Xu, D. (1983) Robustness of maximum likelihood estimates for multi-step predictions: the Exponential Smoothing case, *Biometrika*, 80, 623-641.
- Tiao G.C., Tsay, R.S. (1994) Some advances in non-linear and adaptive modelling in time series, *Journal of Forecasting*, 13, 109-131.
- Walker A.M. (1964) Asymptotic properties of least-squares estimates of parameters of the spectrum of a stationary non-deterministic time series, *Journal of the Australian Mathematical Society*, 4, 363-384.