Transfer Function Models with Time-varying Coefficients

Toloi, Clelia
University of São Paulo, Department of Statistics
São Paulo 05508-090, Brazil
clelia@ime.usp.br

Moura, Maria Sílvia
Federal University of São Carlos, Department of Statistics
São Carlos 13560-000, Brazil
msilvia@power.ufscar.br

Morettin, Pedro
University of São Paulo, Department of Statistics
São Paulo 05508-090, Brazil
pam@ime.usp.br

Introduction

Transfer function models became very popular after the introduction of the so-called Box and Jenkins models, see Box et al. (1994). A natural extension of the transfer function model for the case of nonstationary time series that are both integrated of order d is the model

$$(1) \qquad Y_t = \frac{\omega_0(t) + \omega_1(t)B + \ldots + \omega_r(t)B^r}{1 - \delta_1(t)B - \ldots - \delta_s(t)B^s} X_{t-b} + (1-B)^{-d} \frac{1 - \theta_1(t)B - \ldots - \theta_q(t)B^q}{1 - \phi_1(t)B - \ldots \phi_p(t)B^p} a_t,$$

where b is the delay, a_t is white noise, with mean zero and constant variance, and $\omega_t(B), \delta_t(B), \theta_t(B)$ and $\delta_t(B)$ are now polynomial operators with time-varying coefficients.

Dahlhaus et al. (1999) considered the model

$$(2) \qquad X_{t,T} = \sum_{j=1}^{p} a_j(t/T) X_{t-j,T} + \sigma(t/T)\varepsilon_t,$$

with ε_t i.i.d.$(0,1)$, the functions a_j supported on the interval $[0,1]$ and connected to the underlying series by an appropriate rescaling of time. Under some conditions, (2) has a sequence of solutions $X_{t,T}$ of the form

$$(3) \qquad X_{t,T} = \sum_{\ell=0}^{\infty} \pi_{t,T,\ell} \varepsilon_{t-\ell},$$

with $\sup_{t,T} \sum_{\ell=0}^{\infty} |\pi_{t,T,\ell}| < \infty$. This implies a.s. convergence of the series in (2). See Künsch (1995) for details.

Chiann and Morettin (1999, 2005) considered linear systems of the form

$$(4) \qquad Y_{t,T} = \sum_{j} a_j(t/T) X_{t-j,T} + \sigma(t/T)\varepsilon_t,$$

where ε_t i.i.d.$(0,1)$ and a_j as in (2), satisfying further conditions. Dahlhaus et al. (1999) and Chiann and Morettin (1999, 2005) used wavelet expansions of the time-varying coefficients.

In this paper we consider the model

$$(5)\quad Y_{t,T} = \sum_{i=1}^{s} \delta_i(t) Y_{t-i,T} + \sum_{j=0}^{r} \omega_j(t) X_{t-j,T} + \varepsilon_t, \quad t = 1, \ldots, T,$$

where ε_t is i.i.d. $(0, \sigma^2)$ and assume that the error and the input series are independent. Assumptions on the functions $\omega_j(t)$, $j = 0, 1, 2, \cdots, r$, and $\delta_j(t)$, $j = 1, 2, \cdots, s$, are given in Section 4. As in Dahlhaus et al. (1999) we assume that the functions δ_j and ω_j are supported on the interval $[0, 1]$.

We consider the problem of estimating $\omega_j(t)$, $j = 0, 1, 2, \cdots, r$ and $\delta_j(t)$, $j = 1, 2, \cdots, s$, in time domain, using wavelet expansions. Basic notions on wavelets are given in Section 2. We use least squares to obtain the estimators of the wavelet coefficients. Then the detail coefficients are shrinked before the inverse wavelet transform is applied to obtain the final estimates of $\omega_j(t)$ and $\delta_j(t)$. This results in a nonlinear smoothing procedure. See Section 2 for further details.

Wavelets

In this section we discuss some basic ideas on wavelets. From two basic functions, the *scaling function* $\phi(x)$ and the *wavelet* $\psi(x)$ we define infinite collections of translated and scaled versions, $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$, $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$, $j, k \in \mathbb{Z}$. We assume that $\{\phi_{\ell,k}(\cdot)\}_{k \in Z} \cup \{\psi_{j,k}(\cdot)\}_{j \geq \ell; k \in \mathbb{Z}}$ forms an orthonormal basis of $L_2(\mathbb{R})$, for some coarse scale ℓ. A key point (Daubechies, 1992) is that it is possible to construct compactly supported ϕ and ψ that generate an orthonormal system and have space-frequency localization, which allows parsimonious representations for wide classes of functions in wavelet series.
In some applications the functions involved are defined in a compact interval, such as $[0, 1]$. This will be the case of our functions $\omega_j(t)$ and $\delta_j(t)$ in (5). So it will be necessary to consider an orthonormal system that spans $L_2([0, 1])$. Several solutions were proposed, the most satisfactory one being that by Cohen et al.(1993). Accordingly, for any function $f \in L_2([0, 1])$, we can expand it in an orthogonal series

$$(6)\quad f(x) = \alpha_{0,0}\phi(x) + \sum_{j \geq 0} \sum_{k \in I_j} \beta_{j,k} \psi_{j,k}(x),$$

with the wavelet coefficients given by

$$(7)\quad \alpha_{0,0} = \int f(x)\phi(x)dx, \quad \beta_{j,k} = \int f(x)\psi_{j,k}(x)dx,$$

and where $I_j = \{k : k = 0, \ldots, 2^{j-1} - 1\}$, taking $\ell = 0$.

Often we consider the sum in (6) for a maximum level J,

$$f(x) \simeq \alpha_{00}\phi(x) + \sum_{j=0}^{J-1} \sum_{k \in I_j} \beta_{jk} \psi_{jk}(x).$$

The thresholding technique consists of reducing the noise included in a signal through the application of a threshold to the estimated wavelets coefficients $\hat{\beta}_{jk}$. Some commonly used forms are the soft and hard thresholds, given by

$$\begin{aligned}
\delta^{(s)}(\hat{\beta}_{jk}, \lambda) &= (|\hat{\beta}_{jk}| - \lambda)_+ \mathrm{sgn}(\hat{\beta}_{jk}), \\
\delta^{(h)}(\hat{\beta}_{jk}, \lambda) &= \hat{\beta}_{jk} I(|\hat{\beta}_{jk}| \geq \lambda),
\end{aligned}$$

respectively.

In this paper we will use ordinary wavelets, as in Dahlhaus et al. (1999), since they work well for our purposes. See also a related discussion in Morettin and Chiann (2008). We will make use in particular of Daubechies least asymmetric wavelets in the simulations and application.

Estimators

Consider model (5), with the orders r and s assumed to be known. The idea is to expand $\delta_i(t),\ i = 1,\ldots,s$, and $\omega_i(t),\ i = 0,1,\ldots,r$, in wavelet series

$$\text{(8)} \qquad \delta_i(u) = \alpha_{00}^{(\delta_i)}\phi_{00}(u) + \sum_{j=0}^{J-1}\sum_{k\in I_j}\beta_{jk}^{(\delta_i)}\psi_{jk}(u), \quad i = 1,\ldots,s,$$

$$\text{(9)} \qquad \omega_i(u) = \alpha_{00}^{(\omega_i)}\phi_{00}(u) + \sum_{j=0}^{J-1}\sum_{k\in I_j}\beta_{jk}^{(\omega_i)}\psi_{jk}(u), \quad i = 0,\ldots,r.$$

The empirical wavelet coefficients are obtained minimizing

$$\text{(10)} \qquad \sum_{t=q+1}^{T}\left(Y_{t,T} - \sum_{i=1}^{s}\delta_i(t)Y_{t-i,T} - \sum_{j=0}^{r}\omega_j(t)X_{t-j,T}\right)^2,$$

with $\delta_i(t)$ and $\omega_i(t)$ replaced by (8)-(9), $q = \max(r,s)$. These empirical wavelet coefficients are then modified using a soft threshold and finally we build estimators of $\delta_i(t)$ and $\omega_i(t)$ by applying the inverse wavelet transform to these thresholded coefficients.

For easy of exposition we restrict our attention from now on to the simple model with $s = 1$ and $r = 0$, namely

$$\text{(11)} \qquad Y_{t,T} = \delta_1(t/T)Y_{t-1,T} + \omega_0(t/T)X_{t,T} + \varepsilon_t.$$

So we regress $Y_{t,T}$ on $X_{t,T}$ and $Y_{t-1,T}$, using the expansions (8)-(9) for $\delta_1(t/T)$ and $\omega_0(t/T)$. Let $\Delta = 2^J - 1$. In matrix notation we have

$$\begin{bmatrix} Y_{2,T} \\ Y_{3,T} \\ \vdots \\ Y_{T,T} \end{bmatrix} = \begin{bmatrix} \phi_{00}\left(\frac{2}{T}\right)Y_{1,T} & \psi_{00}\left(\frac{2}{T}\right)Y_{1,T} & \cdots & \psi_{J-1,\Delta}\left(\frac{2}{T}\right)Y_{1,T} \\ \phi_{00}\left(\frac{3}{T}\right)Y_{2,T} & \psi_{00}\left(\frac{3}{T}\right)Y_{2,T} & \cdots & \psi_{J-1,\Delta}\left(\frac{3}{T}\right)Y_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{00}\left(\frac{T}{T}\right)Y_{T-1,T} & \psi_{00}\left(\frac{T}{T}\right)Y_{T-1,T} & \cdots & \psi_{J-1,\Delta}\left(\frac{T}{T}\right)Y_{T-1,T} \end{bmatrix} \begin{bmatrix} \alpha_{00}^{(\delta_1)} \\ \beta_{00}^{\delta_1} \\ \vdots \\ \beta_{J-1,\Delta}^{(\delta_1)} \end{bmatrix} +$$

$$\text{(12)} \qquad + \begin{bmatrix} \phi_{00}\left(\frac{2}{T}\right)X_{2,T} & \psi_{00}\left(\frac{2}{T}\right)X_{2,T} & \cdots & \psi_{J-1,\Delta}\left(\frac{2}{T}\right)X_{2,T} \\ \phi_{00}\left(\frac{3}{T}\right)X_{3,T} & \psi_{00}\left(\frac{3}{T}\right)X_{3,T} & \cdots & \psi_{J-1,\Delta}\left(\frac{3}{T}\right)X_{3,T} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{00}\left(\frac{T}{T}\right)X_{T,T} & \psi_{00}\left(\frac{T}{T}\right)X_{T,T} & \cdots & \psi_{J-1,\Delta}\left(\frac{T}{T}\right)X_{T,T} \end{bmatrix} \begin{bmatrix} \alpha_{00}^{(\omega_0)} \\ \beta_{00}^{(\omega_0)} \\ \vdots \\ \beta_{J-1,\Delta}^{(\omega_0)} \end{bmatrix} + \begin{bmatrix} \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_T \end{bmatrix}.$$

It follows easily that the least squares estimators of the coefficients are then given by

$$\text{(13)} \qquad \begin{bmatrix} \hat{\boldsymbol\beta}^{(\delta_1)} \\ \hat{\boldsymbol\beta}^{(\omega_0)} \end{bmatrix} = \begin{bmatrix} \Psi_Y'\Psi_Y & \Psi_Y'\Psi_X \\ \Psi_X'\Psi_Y & \Psi_X'\Psi_X \end{bmatrix}^{-1} \begin{bmatrix} \Psi_Y'\mathbf{Y} \\ \Psi_X'\mathbf{Y} \end{bmatrix},$$

where $\Psi_X = [\Phi_X \; \Psi_X^{(0)} \; \cdots \; \Psi_X^{(J-1)}]$, $\Psi_Y = [\Phi_Y \; \Psi_Y^{(0)} \; \cdots \; \Psi_Y^{(J-1)}]$ and

$$\mathbf{\Phi}_X = \begin{bmatrix} \phi_{00}\left(\frac{2}{T}\right) X_{2,T} \\ \phi_{00}\left(\frac{3}{T}\right) X_{3,T} \\ \vdots \\ \phi_{00}\left(\frac{T}{T}\right) X_{T,T} \end{bmatrix},$$

and

$$\mathbf{\Psi}_X^{(m)} = \begin{bmatrix} \psi_{m0}\left(\frac{2}{T}\right) X_{2,T} & \psi_{m1}\left(\frac{2}{T}\right) X_{2,T} & \cdots & \psi_{m,2^m-1}\left(\frac{2}{T}\right) X_{2,T} \\ \psi_{m0}\left(\frac{3}{T}\right) X_{3,T} & \psi_{m1}\left(\frac{3}{T}\right) X_{3,T} & \cdots & \psi_{m,2^m-1}\left(\frac{3}{T}\right) X_{3,T} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{m0}\left(\frac{T}{T}\right) X_{T,T} & \psi_{m1}\left(\frac{T}{T}\right) X_{T,T} & \cdots & \psi_{m,2^m-1}\left(\frac{T}{T}\right) X_{T,T} \end{bmatrix}.$$

Having obtained the estimates given by (13) we plug them in (8) and (9), resulting in linear estimates $\hat{\delta}_i(u)$ and $\hat{\omega}_i(u)$. Finally nonlinear smoothed estimators are obtained applying some thresold to the detail coefficients $\hat{\beta}_{jk}^{(i)}$ and these will be denoted by $\tilde{\delta}_i(u)$ and $\tilde{\omega}_i(u)$, respectively.

Properties of empirical coefficients

Now we present some properties of the empirical wavelet coefficients. The techniques used to prove the results are quite evolved and are based on function space theory. Basically we adapt the results of Dahlhaus et al. (1999) for the transfer function model (5).

We assume that functions $\omega_i(u),\ i = 0, 1, \ldots, r$ and $\delta_i(u),\ i = 1, \ldots, s$ belong to some function spaces $\mathcal{F}_i$ given by

$$\mathcal{F}_i = \left\{ f = \alpha_{00}\phi + \sum_{j,k} \beta_{jk}\psi_{jk} \ : \| \alpha_{00} \| \leq C_{i1}, \| \beta_{..} \|_{m,p,q} \leq C_{i2} \right\},$$

where

$$\| \beta_{..} \|_{m,p,q} = \left(\sum_{j\geq 0} \left[2^{jsp} \sum_{k\in I_j} |\beta_{jk}|^p \right]^{q/p} \right)^{1/q},$$

$s = m + 1/2 - 1/p$. Here, m is the smoothness degree, p and q $(1 \leq p, q \leq \infty)$ specify the norm and C_{i1} and C_{i2} are positive constants. For these function spaces, the following result is valid (see Donoho et al., 1995):

$$(14) \quad \sup_{f_i\in\mathcal{F}_i} \left\{ \sum_{j\geq J} \sum_{k} |\beta_{jk}^{(i)}|^2 \right\} = O\left(2^{-2J\tilde{s}_i}\right).$$

where $\tilde{s}_i = m_i + 1/2 - 1/\tilde{p}_i$, with $\tilde{p}_i = min\{p_i, 2\}$.

These classes contain Besov, Hölder and L_2-Sobolev spaces, see, for example Vidakovic (1999), Dahlhaus et al. (1999) and Triebel (1992).

It can be shown, see Donoho et al. (1995), that the loss in (12) by truncating at level J is of order $T^{-2m_i/(2m_i+1)}$, if we choose J such that $2^{J-1} \leq T^{1/2} \leq 2^J$. This is achieved if $\tilde{s}_i > 1$.

Concerning the wavelets, we assume that ϕ and ψ are compactly supported on $[0, 1]$ and have continuous derivatives up to order $r > m$, with $m = \max m_i$. We denote the spectral norm by $\| \cdot \|_2$ and the sup norm by $\| \cdot \|_\infty$.

In order to analyze the statistical behavior of the estimated coefficients it is convenient to assume that we take expansions of these coefficients as linear combinations of functions in the $\tilde{V}_J$ spaces, generated by $\{\phi_{J,1}, \phi_{J,2}, \cdots, \phi_{J,2^J}\}$. With this basis, we can write

$$(15)\quad Y_{t,T} = \sum_{i=1}^{2^J} \zeta_{J,i}^{(\delta_1)} \phi_{J,i}\left(\frac{t}{T}\right) Y_{t-1,T}(t) + \sum_{i=1}^{2^J} \zeta_{J,i}^{(\omega_0)} \phi_{J,i}\left(\frac{t}{T}\right) X_{t,T} + \gamma_{t,T},$$

where

$$(16)\quad \gamma_{t,T} = \sum_{j \geq J} \sum_{k \in I_j} \beta_{jk}^{(\delta_1)} \psi_{jk}\left(\frac{t}{T}\right) Y_{t-1,T} + \sum_{j \geq J} \sum_{k \in I_j} \beta_{jk}^{(\omega_0)} \psi_{jk}\left(\frac{t}{T}\right) X_{t,T} + \varepsilon_t.$$

Equation (15) in matrix form becomes

$$(17)\quad \mathbf{Y} = \left[\begin{array}{ccc} \boldsymbol{\varphi}_Y & \vdots & \boldsymbol{\varphi}_X \end{array} \right] \left[\begin{array}{c} \boldsymbol{\zeta}^{(\delta_1)} \\ \cdots \\ \boldsymbol{\zeta}^{(\omega_0)} \end{array} \right] + \boldsymbol{\gamma}.c$$

The relationship between $\binom{\boldsymbol{\beta}^{(\delta_1)}}{\boldsymbol{\beta}^{(\omega_0)}}$ and $\binom{\boldsymbol{\zeta}^{(\delta_1)}}{\boldsymbol{\zeta}^{(\omega_0)}}$ is

$$\begin{pmatrix} \boldsymbol{\beta}^{(\delta_1)} \\ \boldsymbol{\beta}^{(\omega_0)} \end{pmatrix} = \boldsymbol{\Gamma} \begin{pmatrix} \boldsymbol{\zeta}^{(\delta_1)} \\ \boldsymbol{\zeta}^{(\omega_0)} \end{pmatrix},$$

where the $\boldsymbol{\Gamma}$ is a $(2^{J+1} \times 2^{J+1})$ block diagonal matrix. The matrix $\boldsymbol{\Gamma}$ does the transformation $(\hat{\alpha}_{00}^{(i)}, \hat{\beta}_{00}^{(i)}, \hat{\beta}_{10}^{(i)}, \hat{\beta}_{11}^{(i)}, \cdots, \hat{\beta}_{J-1,0}^{(i)}, \cdots, \hat{\beta}_{J-1,\Delta}^{(i)})' = \boldsymbol{\Gamma}(\hat{\zeta}_{J,1}^{(i)}, \cdots, \hat{\zeta}_{J,2^J}^{(i)})'$, so each coefficient estimate $\hat{\beta}_{jk}^{(i)}$, $i = \delta_1, \omega_0$ in (13) can be written as $\hat{\beta}_{jk}^{(i)} = \Gamma'_{i,jk}\hat{\zeta}$, and $\|\boldsymbol{\Gamma}_{\delta_1,j,k}\|_{L_2} = \|\boldsymbol{\Gamma}_{\omega_0,j,k}\|_{L_2} = 1$.

Equation (17) can be written in the form

$$(18)\quad \mathbf{Y} = \boldsymbol{\Upsilon}\boldsymbol{\zeta} + \boldsymbol{\gamma},$$

with $\boldsymbol{\Upsilon} = \left[\boldsymbol{\varphi}_Y \vdots \boldsymbol{\varphi}_X\right]$, $\boldsymbol{\zeta} = \binom{\boldsymbol{\zeta}^{(\delta_1)}}{\boldsymbol{\zeta}^{(\omega_0)}}$ and the vector $\boldsymbol{\gamma}$ has lines given by (16). The least squares estimator of $\boldsymbol{\zeta}$ is given by

$$(19)\quad \hat{\boldsymbol{\zeta}} = \left(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon}\right)^{-1} \boldsymbol{\Upsilon}'\mathbf{Y}.$$

We notice that the error term $\boldsymbol{\gamma}$ is not independent from the regressors $\boldsymbol{\Upsilon}$, which means that the estimator is biased.

We now state results on the square error and mean square error of the empirical wavelet coefficients. Dahlhaus et al. (1999) prove asymptotic normality of these coefficients, assuming that the k-th order cumulants of ε_t in model (2) are uniformly bounded and the process $X_{t,T}$ has the moving average representation (3). We believe that a similar result can be proved for our estimators, under appropriate conditions on the processes $\varepsilon_t, X_{t,T}$ and $Y_{t,T}$. This will be the subject of further research.

The following assumptions are needed for the proofs of propositions that follow.

(A1) We assume that ϕ and ψ are compactly supported on $[0, 1]$ and have continuous derivatives up to order $r > m$, with $m = \max m_i$.
(A2) In the estimation procedure we have used first a linear estimator, truncating the wavelet expansion at scale 2^J. In order to get functions with the appropriate smoothness, J was chosen such that $\tilde{s}_i > 1$.
(A3) The matrix $\boldsymbol{\Upsilon}$ in (18) satisfies $E||(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}||_{L_2}^{2+\eta} = O(T^{-2-\eta})$, for some $\eta > 0$.

Proposition 1. Under Assumptions (A1)-(A3) we have that $\| \hat{\boldsymbol{\zeta}} - \boldsymbol{\zeta} \|_\infty^2 = O_p\left(2^J T^{-1} \log(T)\right)$.

Proposition 2. If Assumptions (A1)-(A3) hold, then $E\left(\hat{\beta}_{jk}^{(i)} - \beta_{jk}^{(i)}\right)^2 = O\left(T^{-1}\right)$ holds uniformly in i, k and $j < J$.

Further remarks

In this paper we have proposed an estimation procedure for a transfer function model with time-varying coefficients. Basically it is a least squares procedure, with the use of wavelets to expand the function coefficients. Firstly, linear estimators for the time-varying coefficients are obtained, truncating the wavelet expansion at an appropriate scale. Then thresholds are applied to the empirical wavelet coefficients to obtain nonlinear smoothed estimators for the function coefficients. Simulations have shown that this procedure leads to estimators with a good performance. See Moura et al. (2010) for details and for an empirical application. Some statistical properties for the empirical wavelet coefficients were derived. Further studies are needed on the estimation of the variance and asymptotic normality of the empirical wavelet coefficients, on the rate for the risk of the thresholded estimator over the smoothness classes $\mathcal{F}_i$ and on issues related to the identification and diagnostics for these non-stationary models.

References

Box, G. e. P., Jenkins G. M. and Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control.* englewood Cliffs: Prentice-Hall.

Chiann, C. and Morettin, P.A. (1999). Estimation of time-varying linear systems. *Statistical Inference for Stochastic Processes,* **2**, 253-285.

Chiann, C. and Morettin, P.A. (2005). Time-domain estimation of time-varying linear systems. *Journal of Nonparametric Statistics,* **17***, 365–383.*.

Cohen, A., Daubechies, I. and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Applied Computational and Harmonic Analysis*, **1**, 54-81.

Dahlhaus, R., Neumann, M.H. and von Sachs, R. (1999). Nonlinear wavelet estimation of time-varying autoregressive process. *Bernoulli,* **5**, 873-906.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. SIAM: Philadelphia.

Donoho, D.L., Johnstone, I.M., Kerkyacharian, G. and Picard, D. (1995). Wavelet shrinkage: asymptopia? (with discussion) *Journal of the Royal Satistical Society, Series B*, **57**, 301-337.

Knsch, H. (1995). A Note on causal solutions for locally stationary AR-process. Manuscript, eTH Zrich.

Morettin, P.A. and Chiann, C. (2008). Estimation of functional-coefficient autoregressive models by wavelet methods. *Current Development in Theory and Applications of Wavelets*, **1**, 309-340.

Moura, M.S.A., Morettin, P.A. and Toloi, C.M.C. (2010). Transfer function models with time-varying coefficients. TR, Department of Statistics, University of São Paulo.

Triebel, H. (1992). *Theory of Function Spaces II*. Basel: Birhaser Verlag.

Vidakovic, B. (1999). *Statistical Modelling by Wavelets.* New York: Wiley.