Functional autoregressive time series with dependence on derivatives

de Miranda, J. C. S.
University of São Paulo, Institute of Mathematics and Statistics
Rua do Matão
São Paulo 05508-090, Brazil
E-mail: simon@ime.usp.br

We study functional autoregressive time series models that explicitly present dependence on the derivatives of the past observations. The time series observations are assumed to be real valued functions defined on $\mathbb{R}^m$. Equivalently, the data may be seen as hyper-surfaces in $\mathbb{R}^{m+1}$. We assume they are square integrable w.r.t. the Lebesgue measure on $\mathbb{R}^m$. Since, under regularity conditions, models that are linear on the observation and where the dependence on the derivative of the data is linear, like, for example:

$$(1) \qquad \mathcal{X}_{t+1}(p) = \langle \alpha(p,q), \mathcal{X}_t(q) \rangle + \langle \beta(p,q), \nabla\mathcal{X}_t(q)\cdot \mathcal{V}(q) \rangle + \epsilon_{t+1}(p)$$

are reducible to functional (autoregressive time series) linear models without explicit dependence on the derivatives, we will consider models, where the derivative is present, that are non linear with respect to the data or to the derivatives of the data. These models, if they are not degenerated, can not be reduced to functional linear models. We will focus on the non degenerated quadratic models:

$$(2) \qquad \mathcal{X}_{t+1}(p) = \langle \alpha(p,q), \mathcal{X}_t(q)^2 \rangle + \langle \beta(p,q), \nabla\mathcal{X}_t(q)\cdot \mathcal{V}(q) \rangle + \epsilon_{t+1}(p)$$

and

$$(3) \qquad \mathcal{X}_{t+1}(p) = \langle \alpha(p,q), \mathcal{X}_t(q) \rangle + \langle \beta(p,q), \nabla\mathcal{X}_t(q)\cdot \nabla\mathcal{X}_k(q) \rangle + \epsilon_{t+1}(p),$$

since the ideas presented in their analyses are extensible to more general models.

Here p, q are points in $\mathbb{R}^m$, $\alpha : \mathbb{R}^{2m} \to \mathbb{R}$ and $\beta : \mathbb{R}^{2m} \to \mathbb{R}$ are square integrable, $\mathcal{V} : \mathbb{R}^m \to \mathbb{R}^m$ is a known vector field, the sequence $\epsilon_t : \mathbb{R}^m \to \mathbb{R}$, for $t > 0$, is i.i.d. zero mean functional noise, and, for all p, $\beta(p,\cdot\,)$ is assumed to be smooth with compact support. Scalar product of fields is denoted by the symbol $\cdot$ and $\langle \cdot\,,\cdot \rangle$ stands for the inner product w.r.t. the common variable, more precisely, $\langle f(p,q), g(q) \rangle = \int_{\mathbb{R}^m} f(p,q)g(q)dq$.

Given a set of observations $\{\mathcal{X}_t : 1 \le t \le T\}$, we present estimators $\hat{\alpha}$ and $\hat{\beta}$ of α and β that are based on orthonormal basis expansions.

Identifiability

Now we address the identifiability issue. Concerning the (non degenerated) functional time series models (2) or (3), if different pairs of functions (α, β) and (γ, δ) were such that they would generate the same stochastic process, then these models would be unidentifiable. In the case of model (2), non identifiability holds if and only if, there exist different pairs of functions (α, β) and (γ, δ) such that for all $\mathcal{X}$ the following equality:

$$(4) \qquad \langle \alpha(p,q), \mathcal{X}(q)^2 \rangle + \langle \beta(p,q), \nabla\mathcal{X}(q)\cdot \mathcal{V}(q) \rangle = \langle \gamma(p,q), \mathcal{X}(q)^2 \rangle + \langle \delta(p,q), \nabla\mathcal{X}(q)\cdot \mathcal{V}(q) \rangle$$

is fulfilled. However, we can prove that if we have $\langle f(p,q), \mathcal{X}(q)^2 \rangle + \langle g(p,q), \nabla\mathcal{X}(q)\cdot \mathcal{V}(q) \rangle = 0$, for all $\mathcal{X}$, then $f = 0 = g$. Thus, $\alpha(p,q) = \gamma(p,q)$ and $\beta(p,q) = \delta(p,q)$, and the the functional time series model **(2) is identifiable**. Concerning the model (3), the requirement: for all $\mathcal{X}$,

$$(5) \qquad \langle \alpha(p,q), \mathcal{X}_t(q) \rangle + \langle \beta(p,q), \nabla\mathcal{X}_t(q)\cdot \nabla\mathcal{X}_k(q) \rangle = \langle \gamma(p,q), \mathcal{X}_t(q) \rangle + \langle \delta(p,q), \nabla\mathcal{X}_t(q)\cdot \nabla\mathcal{X}_k(q) \rangle,$$

also implies $(\alpha, \beta) = (\gamma, \delta)$ and the the functional time series model **(3) is identifiable**.

Estimation

In order to estimate the functional parameters that appear in models (2) and (3) we will expand them in series using an orthonormal basis. Let $\mathcal{B} = \{\varphi_i : i \in \mathcal{I}\}$ be an orthonormal basis of $L^2(\mathbb{R}^m)$. The set $\{\varphi_i \otimes \varphi_j : i, j \in \mathcal{I}\}$ is an orthonormal basis of $L^2(\mathbb{R}^{2m})$.

We will need the following definitions:

Definition 1: (Product connexion coefficients) *For all i, j in $\mathcal{I}$ let*

$$\varphi_i \varphi_j = \sum_{k \in \mathcal{I}} \Gamma^k_{i,j} \varphi_k. \tag{6}$$

For all i, j and k in $\mathcal{I}$, the uniquely determined coefficients $\Gamma^k_{i,j}$ in (6) will be called product connexion coefficients.

Definition 2: (Differential connexion coefficients) *Let $\{e_j : 1 \le j \le m\}$ be the canonical basis of $\mathbb{R}^m$. For all i in $\mathcal{I}$ let*

$$\nabla \varphi_i = (\sum_{k \in \mathcal{I}} \Delta_i^{1,k} \varphi_k, \ldots, \sum_{k \in \mathcal{I}} \Delta_i^{m,k} \varphi_k) = \sum_{j=1}^{m} \sum_{k \in \mathcal{I}} \Delta_i^{j,k} \varphi_k e_j \tag{7}$$

For all i, k in $\mathcal{I}$, and $j \in \mathbb{N}, 1 \le j \le m$, the uniquely determined coefficients $\Delta_i^{j,k}$ in (7) will be called differential connexion coefficients. In this way, partial derivatives are written as $\partial_j \varphi_i = \sum_{k \in \mathcal{I}} \Delta_i^{j,k} \varphi_k$.

Now, using $\mathcal{B} \otimes \mathcal{B}$, we write

$$\alpha = \sum_{i,j \in \mathcal{I}} \mathsf{a}^{i,j} \varphi_i \otimes \varphi_j \text{ and } \beta = \sum_{i,j \in \mathcal{I}} \mathsf{b}^{i,j} \varphi_i \otimes \varphi_j. \tag{8}$$

Once we find estimators $\hat{\mathsf{a}}^{i,j}$ for $\mathsf{a}^{i,j}$ and $\hat{\mathsf{b}}^{i,j}$ for $\mathsf{b}^{i,j}$ we can construct estimators, $\hat{\alpha}$, and $\hat{\beta}$ of α and β by substituting $\hat{\mathsf{a}}^{i,j}$ for $\mathsf{a}^{i,j}$ and $\hat{\mathsf{b}}^{i,j}$ for $\mathsf{b}^{i,j}$ in (8).

The representations, using bases $\mathcal{B}$ and $\mathcal{B} \otimes \mathcal{B}$, of models (2) and (3) are written as

$$\sum_{i \in \mathcal{I}} \mathsf{x}^i_{t+1} \varphi_i = \langle \sum_{i,j \in \mathcal{I}} \mathsf{a}^{i,j} \varphi_i \otimes \varphi_j, (\sum_{i \in \mathcal{I}} \mathsf{x}^i_t \varphi_i)^2 \rangle +$$

$$\langle \sum_{i,j \in \mathcal{I}} \mathsf{b}^{i,j} \varphi_i \otimes \varphi_j, (\nabla \sum_{i \in \mathcal{I}} \mathsf{x}^i_t \varphi_i) \cdot (\sum_{j=1}^{m} \sum_{i \in \mathcal{I}} \mathsf{v}^{j,i} \varphi_i e_j) \rangle + \sum_{i \in \mathcal{I}} \mathsf{e}^i_{t+1} \varphi_i \tag{9}$$

and

$$\sum_{i \in \mathcal{I}} \mathsf{x}^i_{t+1} \varphi_i = \langle \sum_{i,j \in \mathcal{I}} \mathsf{a}^{i,j} \varphi_i \otimes \varphi_j, \sum_{i \in \mathcal{I}} \mathsf{x}^i_t \varphi_i \rangle +$$

$$\langle \sum_{i,j \in \mathcal{I}} \mathsf{b}^{i,j} \varphi_i \otimes \varphi_j, (\nabla \sum_{i \in \mathcal{I}} \mathsf{x}^i_t \varphi_i) \cdot (\nabla \sum_{i \in \mathcal{I}} \mathsf{x}^i_t \varphi_i) \rangle + \sum_{i \in \mathcal{I}} \mathsf{e}^i_{t+1} \varphi_i, \tag{10}$$

where $\mathcal{X}_t = \sum_{i \in \mathcal{I}} \mathsf{x}^i_t \varphi_i$, $\mathcal{V} = \sum_{j=1}^{m} \sum_{i \in \mathcal{I}} \mathsf{v}^{j,i} \varphi_i e_j$, and $\epsilon_t = \sum_{i \in \mathcal{I}} \mathsf{e}^i_t \varphi_i$.

Using connexion coefficients, equalities (9) and (10) can be written as

$$\forall i \in \mathcal{I} \;\; \mathsf{x}^i_{t+1} = \sum_{j,k,l \in \mathcal{I}} \mathsf{a}^{i,j} \mathsf{x}^k_t \mathsf{x}^l_t \Gamma^j_{k,l} + \sum_{d=1}^{m} \sum_{j,k,l,s \in \mathcal{I}} \mathsf{b}^{i,j} \mathsf{x}^s_t \mathsf{v}^{d,l} \Delta_s^{d,k} \Gamma^j_{k,l} + \mathsf{e}^i_{t+1} \tag{11}$$

and

$$\forall i \in \mathcal{I} \;\; \mathsf{x}^i_{t+1} = \sum_{j \in \mathcal{I}} \mathsf{a}^{i,j} \mathsf{x}^j_t + \sum_{d=1}^{m} \sum_{j,k,l,s,w \in \mathcal{I}} \mathsf{b}^{i,j} \mathsf{x}^w_t \mathsf{x}^s_t \Delta_w^{d,k} \Delta_s^{d,l} \Gamma^j_{k,l} + \mathsf{e}^i_{t+1}. \tag{12}$$

The systems of equations (11) and (12) are the basic relations from which we can construct estimators of $\mathsf{a}^{i,j}$ and $\mathsf{b}^{i,j}$ and, consequently, of α and β. In practice a conveniently chosen finite subset $\mathcal{F}$ of $\mathcal{B}$ will be used in place of $\mathcal{I}$ in the systems (11) and (12). The size of $\mathcal{F}$ will be chosen in accordance with the length of the time series, typically, for asymptotic reasonings, in such a way that the ratio of the size of $\mathcal{F}$ to the time series length, T, goes to zero as both this length and $\#\mathcal{F}$ go to infinity. For practical concerns, $\#\mathcal{F}$ is chosen in such a way that one is able to nicely estimate the approximate coefficients $\tilde{\mathsf{a}}^{i,j}$ and $\tilde{\mathsf{b}}^{i,j}$, using $\hat{\mathsf{a}}^{i,j}$ and $\hat{\mathsf{b}}^{i,j}$, in the approximate systems

$$\forall t, 1 \le t \le T-1 \ \ \forall i \in \mathcal{F} \ \ \mathsf{x}_{t+1}^i = \sum_{j,k,l \in \mathcal{F}} \tilde{\mathsf{a}}^{i,j} \mathsf{x}_t^k \mathsf{x}_t^l \Gamma_{k,l}^j + \sum_{d=1}^{m} \sum_{j,k,l,s \in \mathcal{F}} \tilde{\mathsf{b}}^{i,j} \mathsf{x}_t^s \mathsf{v}^{d,l} \Delta_s^{d,k} \Gamma_{k,l}^j + \mathsf{e}_{t+1}^i \tag{13}$$

and

$$\forall t, 1 \le t \le T-1 \ \ \forall i \in \mathcal{F} \ \ \mathsf{x}_{t+1}^i = \sum_{j \in \mathcal{F}} \tilde{\mathsf{a}}^{i,j} \mathsf{x}_t^j + \sum_{d=1}^{m} \sum_{j,k,l,s,w \in \mathcal{F}} \tilde{\mathsf{b}}^{i,j} \mathsf{x}_t^w \mathsf{x}_t^s \Delta_w^{d,k} \Delta_s^{d,l} \Gamma_{k,l}^j + \mathsf{e}_{t+1}^i. \tag{14}$$

In this article, we will estimate the approximate coefficients using functional least squares, i.e., our estimation will be guided by the criterion of minimization of the sum of the squares of the L^2 distances between the time series observed functional values and the expected ones predicted by the functional time series model, or, equivalently, by the minimization of the total energy of the errors.

Now, form (2) and (3) we have:

$$\sum_{t=1}^{T-1} \|\mathcal{X}_{t+1}(p) - \langle \alpha(p,q), \mathcal{X}_t(q)^2 \rangle - \langle \beta(p,q), \nabla \mathcal{X}_t(q) \cdot \mathcal{V}(q) \rangle\|_2^2 =$$

$$\sum_{t=1}^{T-1} \|\epsilon_{t+1}(p)\|_2^2 = \sum_{t=1}^{T-1} \sum_{i \in \mathcal{I}} (\mathsf{e}_{t+1}^i)^2 \approx \sum_{t=1}^{T-1} \sum_{i \in \mathcal{F}} (\mathsf{e}_{t+1}^i)^2 \tag{15}$$

and

$$\sum_{t=1}^{T-1} \|\mathcal{X}_{t+1}(p) - \langle \alpha(p,q), \mathcal{X}_t(q) \rangle - \langle \beta(p,q), \nabla \mathcal{X}_t(q) \cdot \nabla \mathcal{X}_k(q) \rangle\|_2^2 =$$

$$\sum_{t=1}^{T-1} \|\epsilon_{t+1}(p)\|_2^2 = \sum_{t=1}^{T-1} \sum_{i \in \mathcal{I}} (\mathsf{e}_{t+1}^i)^2 \approx \sum_{t=1}^{T-1} \sum_{i \in \mathcal{F}} (\mathsf{e}_{t+1}^i)^2, \tag{16}$$

and the minimization criterion applied to these models is equivalent to ordinary least squares estimation of the coefficients in (11) and (12) or, approximately, in (13) and (14). The estimators $\hat{\mathsf{a}}^{i,j}$ and $\hat{\mathsf{b}}^{i,j}$ are those that minimize

$$\sum_{t=1}^{T-1} \sum_{i \in \mathcal{F}} \left(\mathsf{x}_{t+1}^i - \sum_{j,k,l \in \mathcal{F}} \tilde{\mathsf{a}}^{i,j} \mathsf{x}_t^k \mathsf{x}_t^l \Gamma_{k,l}^j - \sum_{d=1}^{m} \sum_{j,k,l,s \in \mathcal{F}} \tilde{\mathsf{b}}^{i,j} \mathsf{x}_t^s \mathsf{v}^{d,l} \Delta_s^{d,k} \Gamma_{k,l}^j \right)^2 \tag{17}$$

and

$$\sum_{t=1}^{T-1} \sum_{i \in \mathcal{F}} \left(\mathsf{x}_{t+1}^i - \sum_{j \in \mathcal{F}} \tilde{\mathsf{a}}^{i,j} \mathsf{x}_t^j - \sum_{d=1}^{m} \sum_{j,k,l,s,w \in \mathcal{F}} \tilde{\mathsf{b}}^{i,j} \mathsf{x}_t^w \mathsf{x}_t^s \Delta_w^{d,k} \Delta_s^{d,l} \Gamma_{k,l}^j \right)^2 \tag{18}$$

From here on, to shorten the lengthy equations, we will use the summation convention, i.e., excluding the index t, *if an index appears twice in a monomial then summation in this index is implied.*

Standard minimization techniques lead to the following systems of equations that must be satisfied by the estimators $\hat{\mathsf{a}}^{i,j}$ and $\hat{\mathsf{b}}^{i,j}$. Note that these estimators, a priori, depend on $\mathcal{F}$.

$$\forall (i,j) \in \mathcal{F}^2 \quad \sum_{t=1}^{T-1} \left(\left(\mathsf{x}^i_{t+1} - \hat{\mathsf{a}}^{i,j\prime} \mathsf{x}^k_t \mathsf{x}^l_t \Gamma^{j\prime}_{k,l} - \hat{\mathsf{b}}^{i,j\prime} \mathsf{x}^s_t \mathsf{v}^{d,l} \Delta^{d,k}_s \Gamma^{j\prime}_{k,l} \right) \left(\mathsf{x}^k_t \mathsf{x}^l_t \Gamma^j_{k,l} \right) \right) = 0$$

$$\text{(19)} \qquad \sum_{t=1}^{T-1} \left(\left(\mathsf{x}^i_{t+1} - \hat{\mathsf{a}}^{i,j\prime} \mathsf{x}^k_t \mathsf{x}^l_t \Gamma^{j\prime}_{k,l} - \hat{\mathsf{b}}^{i,j\prime} \mathsf{x}^s_t \mathsf{v}^{d,l} \Delta^{d,k}_s \Gamma^{j\prime}_{k,l} \right) \left(\mathsf{x}^s_t \mathsf{v}^{d,l} \Delta^{d,k}_s \Gamma^j_{k,l} \right) \right) = 0$$

and

$$\forall (i,j) \in \mathcal{F}^2 \quad \sum_{t=1}^{T-1} \left(\mathsf{x}^i_{t+1} - \hat{\mathsf{a}}^{i,j\prime} \mathsf{x}^{j\prime}_t - \hat{\mathsf{b}}^{i,j\prime} \mathsf{x}^w_t \mathsf{x}^s_t \Delta^{d,k}_w \Delta^{d,l}_s \Gamma^{j\prime}_{k,l} \right) (\mathsf{x}^j_t) = 0$$

$$\text{(20)} \qquad \sum_{t=1}^{T-1} \left(\left(\mathsf{x}^i_{t+1} - \hat{\mathsf{a}}^{i,j\prime} \mathsf{x}^{j\prime}_t - \hat{\mathsf{b}}^{i,j\prime} \mathsf{x}^w_t \mathsf{x}^s_t \Delta^{d,k}_w \Delta^{d,l}_s \Gamma^{j\prime}_{k,l} \right) \left(\mathsf{x}^w_t \mathsf{x}^s_t \Delta^{d,k}_w \Delta^{d,l}_s \Gamma^j_{k,l} \right) \right) = 0$$

Observe that both systems of equations (19) and (20) are **linear in** $\hat{\mathsf{a}}^{i,j}$ and $\hat{\mathsf{b}}^{i,j}$, the estimators of the coefficients. This is of great advantage for applications due to computational reasons.

The final estimators are written

$$\text{(21)} \qquad \hat{\alpha} = \hat{\alpha}_{\mathcal{F}} = \sum_{i,j \in \mathcal{F}} \hat{\mathsf{a}}^{i,j} \varphi_i \otimes \varphi_j \text{ and } \hat{\beta} = \hat{\beta}_{\mathcal{F}} = \sum_{i,j \in \mathcal{F}} \hat{\mathsf{b}}^{i,j} \varphi_i \otimes \varphi_j.$$

Stability

Let us now study the stability of the dynamical systems defined by equations (2) and (3). We will establish sufficient conditions on the functional parameters and on the functional noise that will guarantee that the orbits of all functions that belong to a ball centered at the origin of an appropriate function space will be bounded and will all remain inside this ball. Now we will consider the space $\mathcal{C}^1_0(\mathbb{R}^m)$ of continuously differentiable real functions defined on $\mathbb{R}^m$ with compact support. Define, for $f \in \mathcal{C}^1_0(\mathbb{R}^m)$, $\|f\|_{1,\infty} = \|f\|_\infty \vee \left(\bigvee_{i=1}^m \|\nabla^i f\|_\infty\right)$. Denote $Supp f$ the support of a function f. The Lebesgue measure is denoted by ℓ.

Theorem 1: *Consider the model (2). Let for all* $q \in \mathbb{R}^m$, $\alpha(\cdot, q)$ *and* $\beta(\cdot, q)$ *belong to* $\mathcal{C}^1_0(\mathbb{R}^m)$ *and suppose that, for all* $t \in \mathbb{N}^*$ *and a.s., the trajectories of the functional noise belongs to* $\mathcal{C}^1_0(\mathbb{R}^m)$*. Assume also that all innovations are uniformly bounded in the* $\| \ \|_{1,\infty}$ *norm, i.e., there exists* $\eta > 0$ *such that for all* t*, and a.s.,* $\|\epsilon_t\|_{1,\infty} < \eta$*. Let* $A = \sup_p \ell(Supp\ \alpha(p, \cdot)) \sup_q \|\alpha(\cdot, q)\|_{1,\infty}$ *and* $B = m \sup_p \ell(Supp\ \beta(p, \cdot)) \sup_q \|\beta(\cdot, q)\|_{1,\infty} \ (\bigvee_{i=1}^m \|\mathcal{V}^i\|_\infty)$*. Suppose that* $B < 1$ *and* $A \leq (1-B)^2/(4\eta)$*. Then, the following stability result holds: For any* $t_0 \in \mathbb{N}^*$*, if* $\mathcal{X}_{t_0} \in B[0, r] \subset (\mathcal{C}^1_0(\mathbb{R}^m), \| \ \|_{1,\infty})$ *then, for all* $t \geq t_0$*,* $\mathcal{X}_t \in B[0, r]$*, where* $r = \frac{1-B+\sqrt{(1-B)^2-4A\eta}}{2A}$.

Analogously, we have the following:

Theorem 2: *Consider the model (3). Let for all* $q \in \mathbb{R}^m$, $\alpha(\cdot, q)$ *and* $\beta(\cdot, q)$ *belong to* $\mathcal{C}^1_0(\mathbb{R}^m)$ *and suppose that, for all* $t \in \mathbb{N}^*$ *and a.s., the trajectories of the functional noise belongs to* $\mathcal{C}^1_0(\mathbb{R}^m)$*. Assume also that all innovations are uniformly bounded in the* $\| \ \|_{1,\infty}$ *norm, i.e., there exists* $\eta > 0$ *such that for all* t*, and a.s.,* $\|\epsilon_t\|_{1,\infty} < \eta$*. Let* $A = m \sup_p \ell(Supp\ \beta(p, \cdot)) \sup_q \|\beta(\cdot, q)\|_{1,\infty}$ *and* $B = \sup_p \ell(Supp\ \alpha(p, \cdot)) \sup_q \|\alpha(\cdot, q)\|_{1,\infty}$*. Suppose that* $B < 1$ *and* $A \leq (1-B)^2/(4\eta)$*. Then, the following stability result holds: For any* $t_0 \in \mathbb{N}^*$*, if* $\mathcal{X}_{t_0} \in B[0, r] \subset (\mathcal{C}^1_0(\mathbb{R}^m), \| \ \|_{1,\infty})$ *then, for all* $t \geq t_0$*,* $\mathcal{X}_t \in B[0, r]$*, where* $r = \frac{1-B+\sqrt{(1-B)^2-4A\eta}}{2A}$.

Final Remarks

Although we have considered the observations of the time series $\mathcal{X}_t$ to be defined on $\mathbb{R}^m$, we could have considered them to be defined on a subset $\mathcal{O}$ of $\mathbb{R}^m$. In this case we would have chosen $\mathcal{B}$ to be an orthonormal basis of $L^2(\mathcal{O})$. We observe here that the innovations ϵ_t are assumed to be *a.s.* differentiable. The time series models (2) and (3) are among the simplest functional time series models where dependence on the derivatives of the observations is present. However, the idea of writing these derivatives as linear combinations of the derivatives of the basis' functions through the use of derivative connexion coefficients, and the possibility of writing products of functions as linear combinations of basis' functions, made possible by the use of product connexion coefficients, permit us to study and to estimate time series models that belong to a large class of models. This class includes models such as polynomial or rational models with autoregressive and moving average features where higher order derivatives are also present. Limit models such as analytic and meromorphic models involving derivatives of arbitrary order are also in this class. Observe that there is no need to define second or higher derivative connexion coefficients; if we denote $\{e_{j,k}: 1 \leq j, k \leq m\}$ the basis of $\mathbb{R}^m \times \mathbb{R}^m$, we can write the second derivative of a function $\mathcal{X} = \sum_{i\in\mathcal{I}} \mathsf{x}^i \varphi_i$ as $\mathsf{d}^2\mathcal{X} = \mathsf{d}\left(\sum_{j=1}^{m} \sum_{i,v\in\mathcal{I}} \mathsf{x}^i \Delta_i^{j,v} \varphi_v e_j\right) = \sum_{k=1}^{m} \sum_{j=1}^{m} \sum_{i,v,w\in\mathcal{I}} \mathsf{x}^i \Delta_i^{j,v} \Delta_v^{k,w} \varphi_w e_{j,k}$. Higher derivatives are written, using the summation convention, simply as $\mathsf{d}^n\mathcal{X} = \mathsf{x}^i \Delta_i^{j_1,v_1} \Delta_{v_1}^{j_2,v_2} \ldots \Delta_{v_{n-1}}^{j_n,w} \varphi_w e_{j_1,\ldots,j_n}$, and they can be used to obtain systems of algebraic equations that represent the functional time series model with respect to the basis chosen. Clearly, we assume that the basis' functions are at least as many times differentiable as the highest order of differentiation that appears in the functional time series model. The fulfillment of this requirement is associated to the construction of the algebraic representations. Note that it is not necessary for the writing of the functions $\mathcal{X}_t$, ϵ_t and their derivatives of any order. (Infinitely smooth functions can be written as *series* of Haar functions, for example, but this is not the point here.) If $\mathcal{O}$ is a parallelepiped in $\mathbb{R}^m$, then the choice of tensor product Fourier sines and cosines basis presents the desirable property of being formed by functions in the class $\mathcal{C}^\infty(\mathcal{O})$. We observe here that this basis also presents the very desired feature that each basis' function has a small number of non zero differential connexion coefficients associated to it. As a matter of fact there are only m of these coefficients that are different from zero associated to each basis' function. Observe that, in general, i.e., with an arbitrary basis $\mathcal{B}$, this number could be infinite. However, there are situations where the use of other basis, like wavelet basis is more suitable. This is the case when the functions α and β present strong localized frequency behaviour. In this situation the wavelet family must satisfy the differentiability requirements imposed by the higher order of differentiation in the time series model.

Acknowledgements: The author thanks **OLSJC.**

REFERENCES

Besse, P. C. ; Cardot, H. and Stephenson, D.B. : Autoregressive forecasting of some functional climatic variations. Scand. J. Statist. 27 673687, (2000).
Dabo-Niang, S. and Ferraty, F. (Editors): Functional and Operatorial Statistics, Contributions to Statistics, Physica-Verlag , a Springer company, Heidelberg (2008).
de Miranda, J. C. S. : Functional analytic and meromorphic models. Preprint.
Ferraty, F. and Vieu, P. : Nonparametric functional data analysis. Springer-Verlag, New York (2006).
Ramsay, J. O. and Silverman, B. W. : Functional data analysis, 2nd edition. Springer- Verlag, New York (2005).

ABSTRACT

We study $\mathbb{R}^{m+1}$*-hyper-surface time series models with dependence on derivatives of past observations. Estimation using orthonormal series expansions of the functional parameters is presented. Product connexion and differential connexion coefficients are defined and used to reduce the functional models to algebraic systems of equations. Two prototype non-linear models are studied. Minimization of the sum of the squares of the* L^2 *norm of the residuals is shown to be equivalent to the minimization of the sum of squared residuals in the algebraic representation. O.L.S. estimation is applied to the systems of algebraic equations associated to these models and the expressions for the estimators are obtained. Results concerning the stability of these dynamical systems are presented. Extensions to more general settings are discussed.*

Keywords: *Non linear functional time series, Autoregressive, Dependence on derivatives, Orthonormal series expansions, Stability, Dynamical system*