

On the Identifiability of the Functional Convolution Model

Giles Hooker

Abstract

This report details conditions under which the Functional Convolution Model described in Asencio et al. (2013) can be identified from Ordinary Least Squares estimates without either dimension reduction or smoothing penalties. We demonstrate that if the covariate functions are not spanned by the space of solutions to linear differential equations, the functional coefficients in the model are uniquely determined in the Sobolev space of functions with absolutely continuous second derivatives.

Asencio et al. (2013) introduced the *Functional Convolution Model* (FCM) in which for each observation $i = 1, \dots, n$ a functional response $Y_i(t)$ depends on the short-term history of one or more functional covariates $X_{ij}(t)$, $j = 1, \dots, p$ which are measured on the same time domain $t \in [0, T_i]$ along with scalars z_{ik} for $k = 1, \dots, d$. This is expressed mathematically

as

$$y_i(t) = \beta_{00} + \sum_{j=1}^d \beta_{0j} z_{ij} + \sum_{j=1}^p \int_0^{\alpha_j} \beta_j(u) x_{ij}(t-u) du + \epsilon_i(t), \quad i = 1, \dots, n. \quad (1)$$

Here $y_i(t)$ responds to the past α_j time units of x_{ij} via a functional linear model (Ramsay and Silverman, 2005) and this relationship is constant over t . The $\epsilon_i(t)$ are assumed to be mean-zero stochastic processes with stationary covariance. Note that while the $y_i(t)$ and the $x_{ij}(t)$ must share the same time domain, this domain needs not be the same across different observations. We have parameterized the model so that $\beta_j(0)$ represents the instantaneous effect of $x_{ij}(t)$ on the response at time t . $\boldsymbol{\beta}_0 = (\beta_{00}, \beta_{01}, \dots, \beta_{0d})$ represent coefficients for scalar covariates including the intercept.

An estimate of the parameters in the FCM via Ordinary Least Squares (OLS) was proposed, obtaining estimates for $\boldsymbol{\beta}(u) = (\boldsymbol{\beta}_0, \beta_1(u), \dots, \beta_p(u))$ that minimize

$$\text{SSE}(\boldsymbol{\beta}) = \sum_{i=1}^n \int_{\alpha^*}^{T_i} \left(y_i(t) - \sum_{j=1}^d \beta_{0j} z_{ij} - \sum_{j=1}^p \int_{t-\alpha_j}^t \beta_j(u) x_{ij}(t-u) du \right)^2 dt$$

for $\alpha^* = \max(\alpha_1, \dots, \alpha_p)$. This has been chosen so that the range of t in the outer integral ensures that the range of $t-u$ within the squared term does not go below 0. To this criterion, Asencio et al. (2013) added smoothing penalties for each of the $\beta_j(u)$ and represented them via a basis expansion.

The identifiability of the FCM under the OLS without penalization, and with arbitrarily complex basis expansions, is not clear. The model can be placed between the scalar

response model $y_i = \beta_0 + \int \beta_1(t)x_i(t)dt + \epsilon_i$ in which $\beta_1(t)$ cannot be identified without smoothing or dimension reduction and the concurrently linear model for functional responses $y_i(t) = \beta_0(t) + \beta_1(t)x_i(t) + \epsilon_i(t)$ in which $\beta_0(t)$ and $\beta_1(t)$ can be obtained from a linear regression at each time t (although smoothing can still serve to reduce variance). The purpose of this report is to investigate under what circumstances both smoothing penalty and basis expansion can be removed. That is, restricting the $\beta_j(u)$ to lie in a Sobolev class of functions, under what conditions does $SSE(\beta)$ have a unique minimum without further restrictions? In general, the design of covariate functions in functional data analysis has received little attention. The functional convolution models studied here present a challenge in that identifiability depends on the finite-sample design of the covariates.

For this report, we assume our covariate processes $\beta_j(t)$ lie in the Sobolev space $W[0, \alpha_j]$ of functions defined on $[0, \alpha_j]$ for which all second derivatives are absolutely continuous. For convenience, we also assume that the $y_i(t)$ and $x_i(t)$ have been centered by their integral averaged over all observations and ignore β_0 . We also assume that the Y_i have been centered by their time-series mean and no other scalar covariates are present in order to remove the intercept from the model. Within this space, minimizing $SSE(\beta)$ is equivalent to setting its Gateaux derivative to zero; that is, solving the variational problem

$$\langle \gamma, G[\beta] \rangle = F(\gamma), \quad \forall \gamma = (\gamma_1, \dots, \gamma_p) \in W[0, \alpha_1] \otimes \dots \otimes W[0, \alpha_p] \quad (2)$$

where

$$F(\gamma) = \sum_{j=1}^p \sum_{i=1}^n \int_0^{\alpha_j} \gamma_j(u) \int_{\alpha^*}^{T_i} x_{ij}(t-u)y_i(t)dtdu,$$

and

$$G[\beta] = \sum_{j,k=1}^p \sum_{i=1}^n \int_0^{\alpha_j} \int_{\alpha^*}^{T_i} x_{ik}(t-v)x_{ij}(t-u)dt\beta_j(u)du.$$

where the inner product is taken as the product L^2 inner product on square-integrable functions

$$\langle \gamma, \beta \rangle = \sum_{j=1}^p \int_0^{\alpha_j} \gamma_j(u)\beta_j(u)du.$$

The identification of the OLS estimates is now equivalent to the invertibility of G . In particular, β will be uniquely identified if

$$\langle \gamma, G[\beta] \rangle = 0, \quad \forall \gamma \Rightarrow \beta = 0$$

and in particular if

$$\langle \beta, G[\beta] \rangle = \int_{\alpha^*}^{T_j} \sum_{i=1}^n \left[\int_0^{\alpha_j} x_{ij}(t-v)\beta_j(v)dv \right] \left[\int_0^{\alpha_j} x_{ij}(t-u)\beta_j(u)du \right] dt > 0 \quad (3)$$

for all j and all non-zero β_j .

In particular, we can let ξ_1, ξ_2, \dots form a basis for $W[0, \alpha_1] \otimes \dots \otimes W[0, \alpha_p]$ with $\xi_j = (\xi_{j1}, \dots, \xi_{jp})$ and (3) reduces to the requirement that for every j and l ,

$$\int_0^{\alpha_j} \xi_{jl}(u)x_{ij}(t-u)du \neq 0 \quad (4)$$

for t in a subset of $[\alpha^*, T_i]$ of positive measure for at least one i . This condition is not readily checked, particularly in real-world applications since it requires checking an infinite collection of inner-products; however the non-identifiability of a design can be readily assessed. Because of this Asencio et al. (2013) employed smoothing parameters to ensure the identifiability of their estimates. However, it is possible to characterize designs such that the collection

$$x_{ijt}(u) = x_{ij}(t - u)$$

spans a finite-dimensional space as t is varied; i.e. for which there is a finite collection of functions $\eta_1(u), \dots, \eta_K(u)$ such that

$$x_{ijt}(u) = \sum_{k=1}^K c_k(t) \eta_k(u)$$

for all t . This set of self-similar functions can be expressed as solutions to a linear differential equation. In the lemma below we restrict to a single real-valued function for the sake of clarity and set $\alpha_j = 1$

Lemma 1. *Let $x(t)$ have continuous first derivatives, then x satisfies*

$$x(t + u) = \sum_{k=1}^K \zeta_k(t) \eta_k(u) \tag{5}$$

for all t and $\eta_k : [0, 1] \rightarrow \mathbb{R}$, if and only if

$$x(t) = \sum_{k=1}^K c_k t^{m_k} e^{a_k t} \sin(b_k t + d_k) \tag{6}$$

for real-valued constants (a_k, b_k, c_k, d_k) , and integers m_k , $k = 1, \dots, K$.

Proof. We observe that

$$x(t + u + dt) = \sum_{k=1}^K \zeta_k(t + dt)\eta_k(u) = \sum_{k=1}^K \zeta_k(t)\eta_k(u + dt)$$

and thus

$$x'(t + u) = \sum_{k=1}^K \zeta_k'(t)\eta_k(u) = \sum_{k=1}^K \zeta_k(t)\eta_k'(u) \quad (7)$$

where $x'(\cdot)$ is the derivative taken with respect to its argument. We can now examine which ζ_k satisfy the second equality in (7) by restricting to a finite set of values for u . Defining a set of evaluation points $u_l = (l - 1)/K$ for $k = 1, \dots, K$ we can produce matrices

$$X_{lk} = \eta_k(u_l), \quad \dot{X} = \eta_k'(u_l).$$

Note that these do not depend on t . The last equality in (7) restricted to u_1, \dots, u_K now defines the differential equation

$$\frac{d}{dt}\boldsymbol{\eta} = \left[X^{-} \dot{X} \right] \boldsymbol{\eta} \quad (8)$$

where X^{-} is a generalized inverse. Solutions to (8) have general form of (6) (e.g., Borelli and Coleman, 2004), and thus $x(t)$ must at least be of this form. It is easy to check that any function of the form (6) satisfies (5) for some K and $(\zeta_k, \eta_k, k = 1, \dots, K)$, completing the converse implication. □

From this we directly obtain the following

Corollary 1. *The variational problem (2) has a unique solution in $W[0, \alpha_1] \otimes \cdots \otimes W[0, \alpha_p]$ if and only if*

$$\sum_{i=1}^n \int_{\alpha^*}^{T_i} \left(x_{ij}(t) - \sum_{k=1}^{K_{ij}} c_{ijk} t^{m_{ijk}} e^{a_{ijk}t} \sin(b_{ijk}t + d_{ijk}) \right)^2 dt > 0$$

for each j and any finite choice of K_{ij} , a_{ijk} , b_{ijk} , c_{ijk} , d_{ijk} and m_{ijk} .

It is not difficult to provide example designs that satisfy (4). Consider the basis of periodic functions on $[0, 1]$ given by $\xi_k(u) = \sin(2k\pi t)$ then taking $x(t)$ to be defined on $[0, 1]$, say, with

$$x(t) = \sum 2^{-k} \xi_k(t)$$

has non-zero inner product on $[0, 1]$ with $\xi_k(u)$ for each k . We note that the range of $x(t)$ need not be restricted to $[0, 1]$. However, between the finite-dimensional design described in Lemma 1 and the identifiable design, it is possible to find $x_i(t)$ that is orthogonal to an infinite dimensional subspace. Continuing our example, setting

$$\tilde{x} = \sum 2^{-4k} \xi_{2k}(t)$$

will yield $\langle \tilde{x}, \xi_k \rangle = 0$ for all k odd provided the domain of x is of integer length. Evaluating identifiability based on finite-dimensional approximations as given, for example in the appendix of Asencio et al. (2013) is straightforward. However, it seems more challenging to provide a protocol for designing experiments for which this model will be employed.

Beyond conditions for identifiability, the implication of these calculations for the convergence rates of the FCM remain un-investigated. As an alternative to the OLS formulation, we can down-sample the data in the FCM to consider only the values of $y_i(t)$ at U intervals. That is we observe that $y_i(lU)$ for $l = 1, \dots, n_i = T_i/U$ taken at the discrete times lU exactly follow a functional linear model. Hence as either $T_i \rightarrow \infty$ or $n \rightarrow \infty$, results in Zhang and Yu (2008) provide the consistency of our estimates when combined with smoothing penalties. Although such penalties needs not be required for the FCM, they were employed in Asencio et al. (2013) and we speculate that in fact the convergence rates that can be obtained for the FCM are no better than the functional linear model.

References

- Asencio, M., G. Hooker, and H. O. Gao (2013). Functional convolution models. *Statistical Modeling to appear*.
- Borelli, R. L. and C. S. Coleman (2004). *Differential Equations: A Modeling Perspective*. New York: John Wiley & Sons.
- Ramsay, J. O. and B. W. Silverman (2005). *Functional Data Analysis*. New York: Springer.
- Zhang, C. and T. Yu (2008). Semiparametric detection of significant activation for brain fMRI. *Annals of Statistics* 36, 1693-1725.