IWFOS'2008

Estimation of functional regression models for functional responses by wavelet approximation

Ana M. Aguilera, Francisco A. Ocaña, Mariano J. Valderrama (*)
Department of Statistics and Operations Research, University of Granada, Spain

Objective

Study some estimation procedures for a functional regression model where both predictor and response variables are functions.

Functional linear model for a functional response

Let us consider a functional predictor $\{X_w : w \in \Omega\} \subset L^2(T)$ and a functional response $\{Y_w\} \subset L^2(S)$, where $(\Omega, \mathcal{A}, \mathcal{P})$ is a probability space, T and S are intervals in $\mathbb{R}$, and both processes are centered.

The sample: $\{(x_w, y_w), w = 1, ..., n\} \subset L^2(T) \times L^2(S)$

The model:

$$\mathrm{E}[Y(s)/x_w] = \int_T \beta(t,s) x_w(t) dt, \quad s \in S. \qquad (1)$$

with $\beta \in L^2(S \times T)$.

The ill–posed **problem:** estimate the β function.

Model estimation

Assuming that X and Y belong to finite dimension spaces spanned by two basis $\{\vartheta_p : p = 1, \ldots, P\}$ and $\{\varphi_q : q = 1, \ldots, Q\}$,

$$x_w(t) = \sum_{p=1}^{P} a_{wp} \vartheta_p(t) \qquad y_w(s) = \sum_{q=1}^{Q} b_{wq} \varphi_q(s),$$

the parameter function $\rightarrow \beta(t,s) = \sum_{p=1}^{P}\sum_{q=1}^{Q} \beta_{pq}\, \vartheta_p(t)\, \varphi_q(s)$.

Model (1) can be formulated as the multivariate linear model

$$\mathbf{B} = \mathbf{A}\Psi\beta + \Upsilon,$$

$\mathbf{B} = (b_{wq})_{n \times Q}$, $\mathbf{A} = (a_{wp})_{n \times P}$, $\Psi = \left(< \vartheta_p, \vartheta_{p'} >_{L^2(T)}\right)_{P \times P}$,
Υ a noise matrix.

Least squares estimation: $\hat{\beta} = ((A\Psi)'(A\Psi))^{-1}(A\Psi)'B$

Problems: multicollinearity and high dimension

A solution: Estimation based on the FPCAs of predictor and response

$$x_w(t) = \sum_{i=1}^{n-1} \xi_{wi}\, f_i(t) \qquad y_w(s) = \sum_{j=1}^{n-1} \eta_{wj}\, g_j(s),$$

where ξ_i and η_j are the PCs of predictor and response curves,

$$\xi_{wi} = \int_T x_w(t)\, f_i(t)\, dt \qquad \eta_{wj} = \int_S y_w(s)\, g_j(s)\, ds\,,$$

with $f_i(t)$ and $g_j(s)$ being their associated PC weights (eigenfunctions of the sample covariance operators).

Model (1) $\Rightarrow$ Linear regression of each PC of Y on all PCs of X

$$\eta_{wj} = \sum_i \xi_{wi}\, \nu_{ij} + \epsilon_{wj} \Rightarrow \beta(t,s) = \sum_{i,j} \nu_{ij}\, f_i(t)\, g_j(s)$$

A functional PC estimation of β can be obtained by

- selecting an optimum set J of PCs of Y
- regressing each of them in terms of and optimum set I_j of PCs of X.

Idea: $R^2 = \dfrac{\mathrm{E}[\|\widehat{Y}\|^2]}{\mathrm{E}[\|Y\|^2]} = \sum_{i,j} P(j,i),$
where $P(j,i)$ is the variance explained by (η_j, ξ_i)
$\Rightarrow P(*,*)$ establishes a priority order in the set of PC pairs $\Rightarrow$

1. Are all the PC pairs needed for estimating β?
   - **Method a**: all possible pairs are considered.
   - **Method s**: pairs with clearly non–significant correlation are leaved out.
2. How many PC pairs?
   - CV (leaving–one–out), BIC, Cp and MSE are adapted to functions (errors $\longrightarrow$ normed errors)
   - Only for simulation studies, $\mathrm{bE} = \|\beta - \widehat{\beta}\|^2_{L^2}$

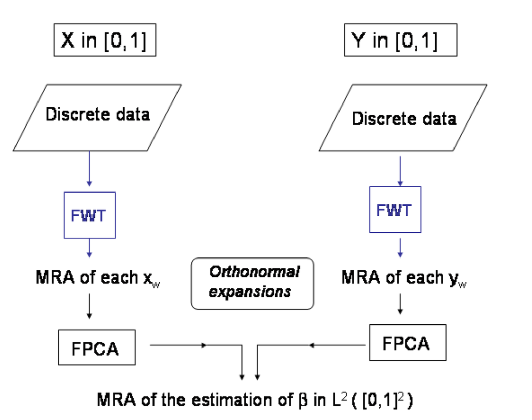
The response $y^*(s)$ associated to a new predictor curve x^* is forecasted

$$y^*(s) = \sum_{j=1}^{J} \eta_j^*\, g_j(s) = \sum_{j=1}^{J}\sum_{i \in I_j} \frac{\sigma_{ij}}{\sigma_i^2}\, \xi_i^*\, g_j(s),$$

where $\xi_i^* = \int_T x^*(t) f_i(t)\, dt$.

Wavelet approximation of sample curves

In practice, basis coefficients of predictor and response sample curves need to be estimated from discrete time observations $\Rightarrow$ Wavelet Analysis



A simulation study

Sketch for each trial

- Predictor process (based on James, Hastie and Sugar (2000)):

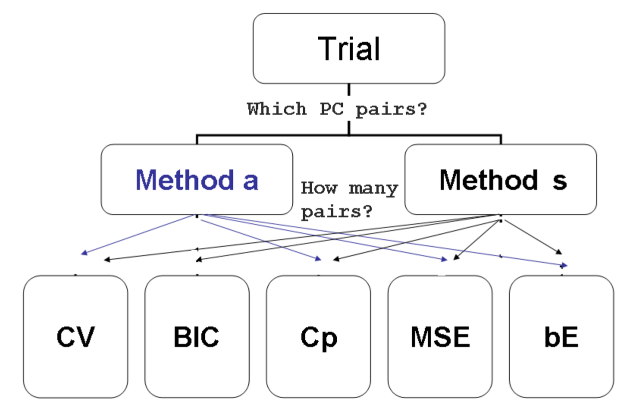
$$x_w(t) = \sum_{p=1}^{14} a_{wp}\, \vartheta_p(t) + \gamma_w, \quad \forall\, t \in [0,1],\ w = 1, \ldots, n = 10,$$

$a_p \leadsto \mathcal{N}(0, |10 - p|)$, $\gamma \leadsto \mathcal{N}(0,1)$,
$\vartheta_{2r-1}(t) = \sin(2\pi rt)$, $\vartheta_{2r}(t) = \cos(2\pi rt)$, $\quad r = 1, \ldots, 7$.
✓ **Discrete data:** evaluate x_w at $t_i = i/20$, $\quad i = 0, \ldots, 20$

- Parameter function: $\beta(s,t) = s\, \sin(2\pi\, t) + \cos(4\,\pi\, t)\,,\ \forall\, s,t \in [0,1]$

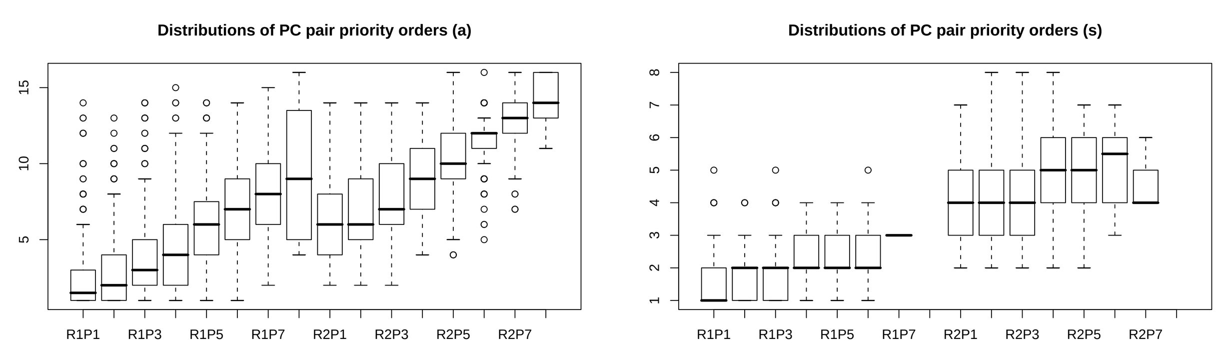
- Response process: $y_w(s) = \int_S \beta(t,s)\, x_w(t)\, dt,\ s \in [0,1]\ w = 1, \ldots, 10$
✓ **Discrete data:** evaluate y_w at $s_j = j/16$, $\quad j = 0, \ldots, 16$

- Estimations of β: (2 methods for considering PC pairs $\times$ 5 criteria for selecting the number of PC pairs)



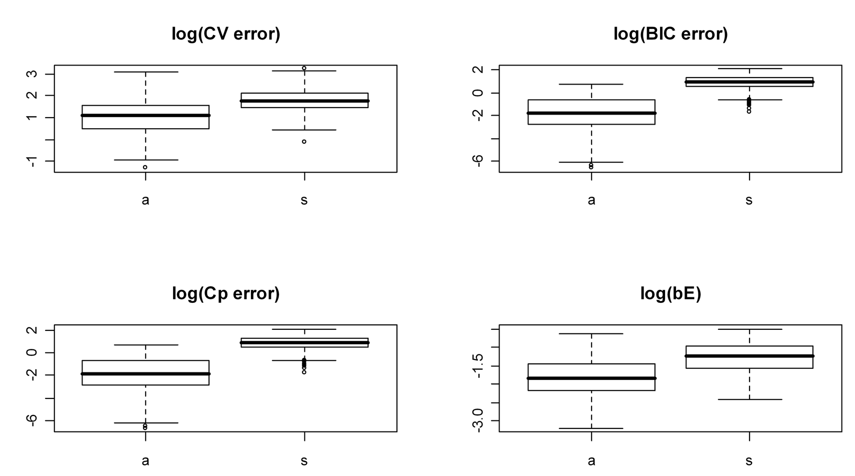
Numerical results (`number of trials = 400`):

Priority orders of PC pairs to be considered to estimate β for Methods **a** and **s**, respectively, where `RjPi` stands for the pair (η_j, ψ_i)

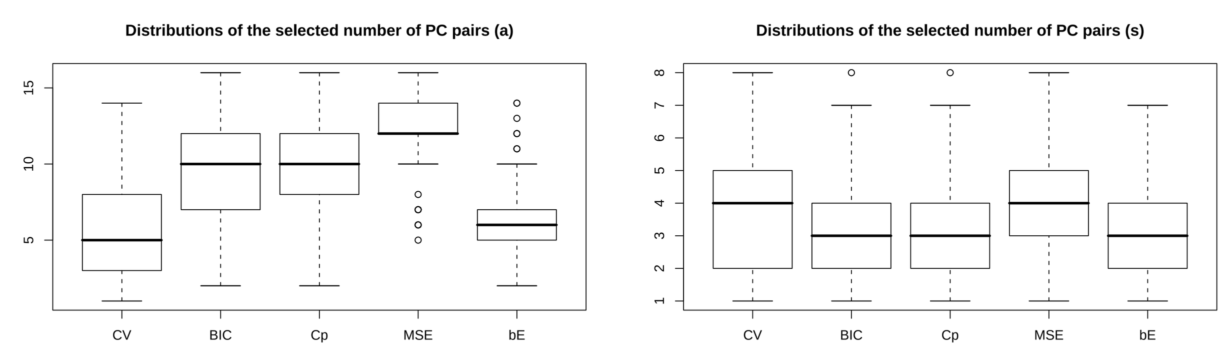


- Study of the models selected by the five considered criteria, being `bE` as the *ideal* reference

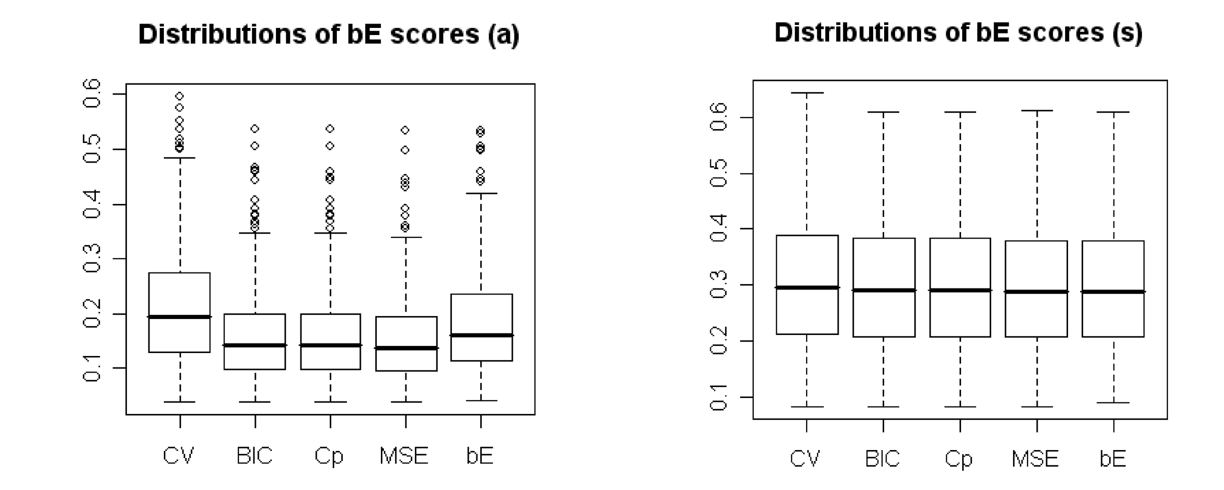
Comparisons of the logarithms of the errors between both methods, for every criterion but MSE

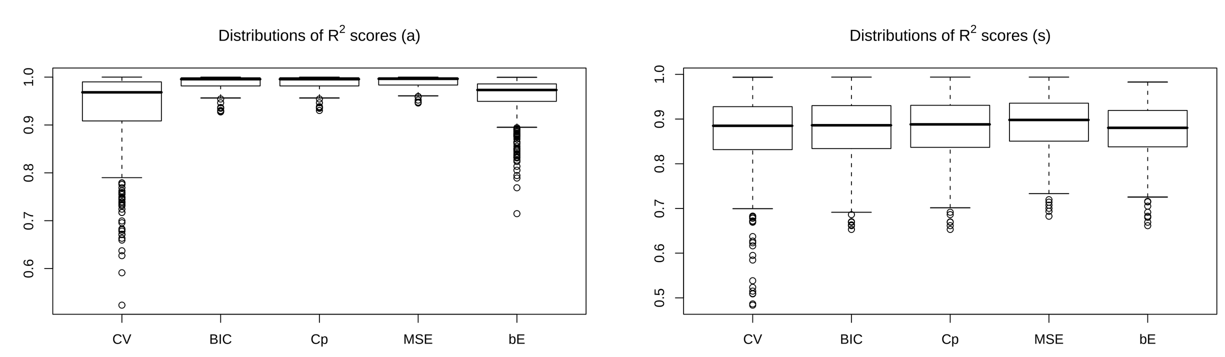


Parsimony of the identified models (# pairs = # parameters)

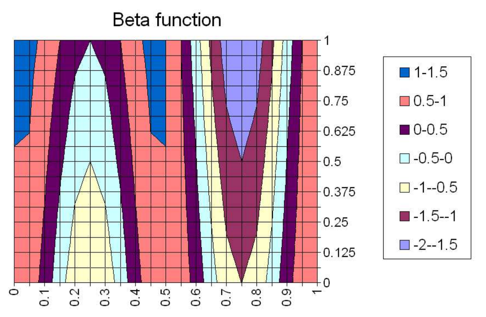


Goodness–of–fit of the models identified by every criterion by using bE

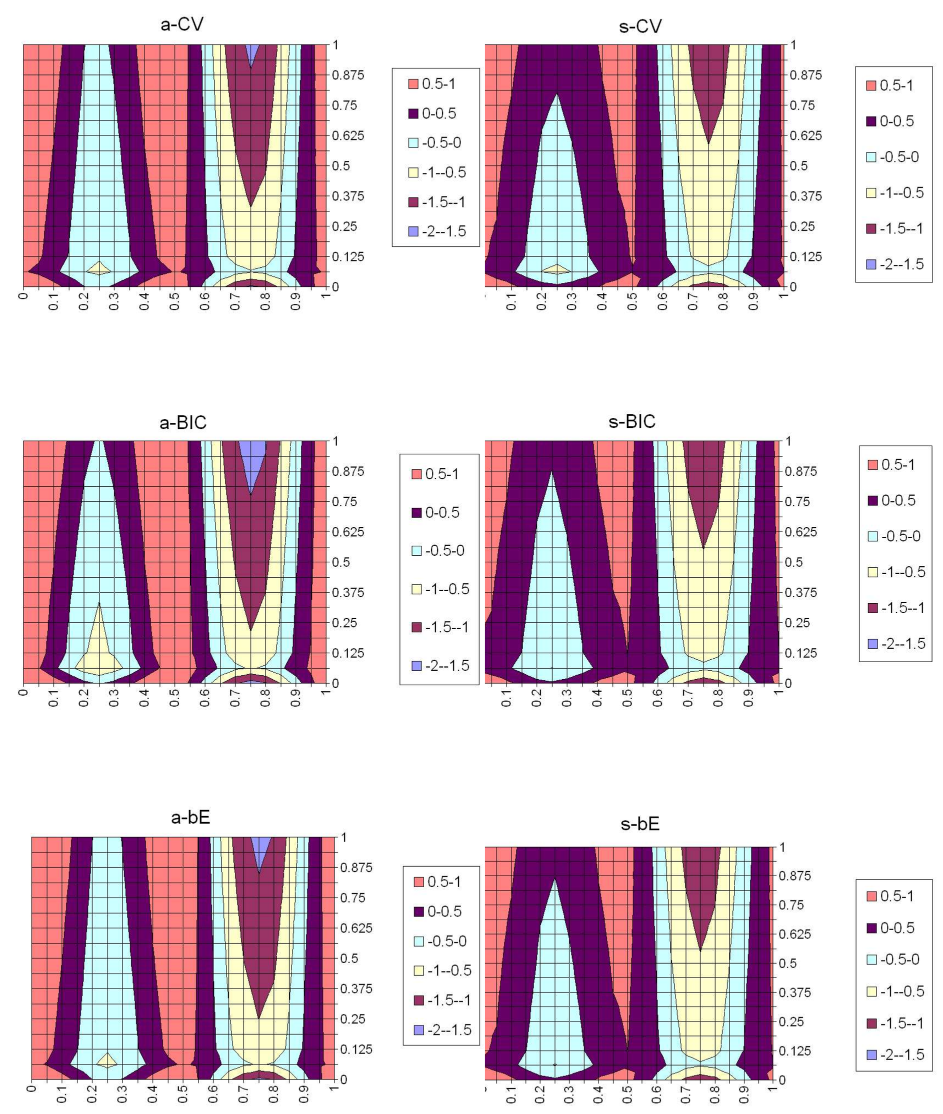




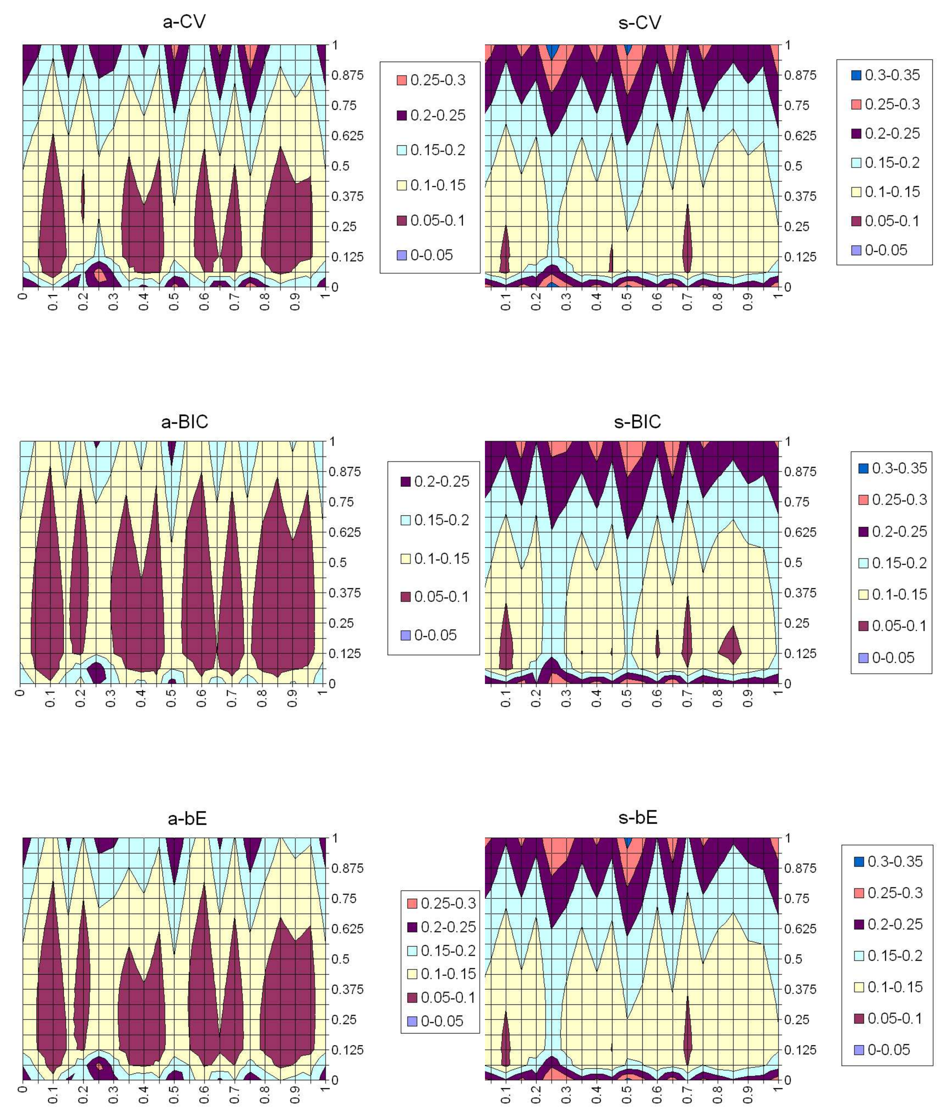
Contour maps of some summary statistic functions



Averages of beta estimations



Variances of beta estimations



Conclusions

1. The priority order established by $P(*,*)$, which lets apply the considered criteria, exhibits a good estimation performance, such as is shown by both methods.
2. Method s provides estimations of β much more parsimonious than Method–a at a non–excessive error cost.
3. Taking into the computational simplicity of BIC and Cp, they would be a good choice for Method s.

Bibliography

- Aguilera, A.M., Ocaña, F.A. and Valderrama, M.J. 1999. Forecasting with unequally spaced data by a functional principal component approach. *Test*, **8**(1), 233-254.
- Ocaña, F.A., Aguilera, A.M. and Escabias, M. 2007. Computational considerations in functional principal component analysis. *Computational Statistics*, **22**, 449-466.
- Cardot, H., Ferraty, F. and Sarda, P. 1999. Functional linear model. *Statistics and Probability Letters*, **45**, 11-22.
- Mallat, S. 1998. *A wavelet tour of signal processing*. Academic Press: San Diego.
- Yao, F., Müller, H-G. and Wang, J-L. 2005. Functional linear regression analysis for longitudinal data. *The Annals of Statistics*, **33**(6), 2873-2903.

(*) Research group URL
`http://www.ugr.es/~predin`