METMA V International Workshop on Spatio-Temporal Modelling
Santiago de Compostela
30 Jun - 2 Jul 2010

Wavelet methods for the functional statistical analysis of fMRI data

R. Fernández-Pascual [1] and M.D. Ruiz-Medina [2,*]

[1] *Department of Statistics and Operation Research, University of Jaén, Campus Las Lagunillas, 23071 Jaén, Spain; rpascual@ujaen.es*
[2] *Department of Statistics and Operation Research, University of Granada, Facultad de Ciencias, Granada, Spain; mruiz@ugr.es*
[*] *Corresponding author*

Abstract. *Nowadays, statistical functional Magnetic Resonance Imaging (fMRI) data analysis emerges as an important area of current research in Neurosciences. Different statistical techniques have been applied for inference purposes, e.g., statistical denoising, reconstruction, parameter estimation, etc. Specifically, in this paper, a functional framework is considered in the derivation of filtering estimators and confidence intervals, in the wavelet domain. A real data example is considered to illustrate results.*

Keywords. *Filtering problem; fMRI data; functional confidence intervals; wavelets*

1 Motivation

The main objective of FMRI studies is to detect changes in areas of the brain, reflected in terms of neuronal activity modifications with respect a general activity pattern. Different statistical methodologies such as general linear models (GLM), multivariate methods (e.g., Principal Component Analysis), Bayesian approaches and statistical shrinkage estimation are studied (see, for example, [2], [6], [3], [4] and [7]), to solve the problem of detection and prediction of brain areas with significative activity from fMRI, positron emission tomography (PET) and magnetoencephalography (MEG) data.

In this paper, we first include a brief introduction to statistical analysis of fMRI data. Second, the orthogonal decomposition of the observed images, in terms of dual Riesz bases, is considered, to approximate the functional least-squares estimators of the elements of the original 2D-image sequence. Asymptotic functional confidence intervals are then derived in the wavelet domain.

2 Statistical models in fMRI

The statistical analysis of fMRI data involves working with massive data sets that present a complicated spatiotemporal structure. The component of interest in a fMRI signal is the BOLD response signal (Blood Oxigentation Level Dependent = ratio of the oxigenated to deoxigenated hemoglobin in the blood, that is, oxygen consumption of active neurons), which is modeled as the convolution of the stimulus function $v(t)$ with the HRF (Hemodinamic Response Function) $h(t)$, supposed to be constant across all voxels, i.e.,

$$x(t) = (v * h)(t).$$

The GLM is given by

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \tag{1}$$

where $\mathbf{Y}$ represent the observed response variable in terms of a linear combination of explanatory variables $\mathbf{X}$ (design matrix), and an error term ε i.i.d. $\sim \mathrm{N}(0,\sigma)$ (see [1]). The parameter β is estimated by applying least-squares method, that is, β is approximated by $\hat{\beta}$, defined as the solution to the following equation system:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}. \tag{2}$$

As fMRI data are autocorrelated, an extension of the above model is then performed by estimating the autocorrelation in the residuals, after model fitting, and removing the autocorrelation by pre-whitening considering a new design matrix $\mathbf{W}^{1/2}\mathbf{X}$ and whitened data $\mathbf{W}^{1/2}\mathbf{y}$, then:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{Y}. \tag{3}$$

Other possible extensions consist on considering that the HRF may vary from voxel to voxel and this structure has been incorporated in the GLM. The idea of different HRFs in different brain regions is modeled by the notion of spatiotemporal basis function (waveforms) (see [1]).

3 Functional estimation

In a functional data context (see [5]),

$$\mathbf{Y}(\cdot) = \mathbf{X}(\cdot) + \varepsilon(\cdot),$$

the Orthogonal Projection Theorem is formulated in terms of Hilbert-valued random variables. Specifically, the functional linear least-squares estimator of $\mathbf{X}$ is given by

$$\hat{\mathbf{X}}(t,\mathbf{x}) = \mathcal{L}_{t,\mathbf{x}}\mathbf{Y}(t,\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^2,\ t \in \mathbb{R}_+,$$

where operator $\mathcal{L}_{t,\mathbf{x}}$ satisfies the following equation:

$$R_{\mathbf{XY}} = \mathcal{L}_{t,\mathbf{x}}R_{\mathbf{YY}}, \tag{4}$$

with $R_{\mathbf{XY}} = E[\mathbf{X}(\cdot) \otimes \mathbf{Y}(\cdot)]$, and $R_{\mathbf{YY}} = E[\mathbf{Y}(\cdot) \otimes \mathbf{Y}(\cdot)]$, and being $\mathbf{Z}(\cdot) \otimes \mathbf{Y}(\cdot)$ the operator given by the tensorial product of the random function $\mathbf{Z}(\cdot)$ with $\mathbf{Y}(\cdot)$.

That is,

$$\hat{\mathbf{X}}(t,\mathbf{x}) = \mathcal{L}_{t,\mathbf{x}}\mathbf{Y}(t,\mathbf{x}) = [R_{\mathbf{XY}}R_{\mathbf{YY}}^{-1}\mathbf{Y}](t,\mathbf{x}). \tag{5}$$

The above equation is approximated here by considering the following orthogonal decomposition of the functional data process $\mathbf{Y}$:

$$\mathbf{Y}(\cdot) \underset{H(\mathbf{Y})}{=} \sum_{\mathbf{k}\in\Gamma_0} \mathbf{Y}(\varphi^{\mathbf{k}})\varphi_{\mathbf{k}}(\cdot) + \sum_{j\geq 0}\sum_{\boldsymbol{\theta}\in\Theta_j} \mathbf{Y}(\gamma^{j:\boldsymbol{\theta}})\gamma_{j:\boldsymbol{\theta}}(\cdot), \tag{6}$$

where $\varphi_{\mathbf{k}} = \mathcal{T}'(\phi_{\mathbf{k}})$, for all $\mathbf{k} \in \Gamma_0$, and $\gamma_{j:\boldsymbol{\theta}} = \mathcal{T}\prime(\psi_{j:\boldsymbol{\theta}})$, for all $\boldsymbol{\theta} \in \Theta_j$, $j \geq 0$. The covariance operator of $\mathbf{Y}$, $R_{\mathbf{YY}}$, is assumed to be factorized as $R_{\mathbf{YY}} = \mathcal{T}'\mathcal{T}$. The system $\left\{\varphi^{\mathbf{k}} : \mathbf{k}\in\Gamma_0\right\} \cup \left\{\gamma^{j:\boldsymbol{\theta}} : \boldsymbol{\theta}\in\Theta_j,\ j\geq 0\right\}$ defines the dual Riesz basis with respect to $L^2(S)$ of the Riesz basis $\left\{\varphi_{\mathbf{k}} : \mathbf{k}\in\Gamma_0^S\right\} \cup \left\{\gamma_{j:\boldsymbol{\theta}} : \boldsymbol{\theta}\in\Theta_j^S,\ j\geq 0\right\}$, constructed from the orthogonal wavelet basis $\left\{\phi_{\mathbf{k}} : \mathbf{k}\in\Gamma_0^S\right\} \cup \left\{\psi_{j:\boldsymbol{\theta}} : \boldsymbol{\theta}\in\Theta_j^S,\ j\geq 0\right\}$ of $L^2(S)$. Here, A' denotes the adjoint operator of A.

3.1 Functional confidence intervals

Under the assumption that our observed functional image sequence is stationary in time, asymptotic functional confidence intervals are derived, from the trace of the empirical wavelet covariance operator of the functional estimate $\hat{\mathbf{X}}$ of $\mathbf{X}$. Specifically, a suitable wavelet basis is selected in order to obtain a weak-dependence model for the functional estimate $\hat{\mathbf{X}}$ in the wavelet domain. Thus, strongly mixing conditions are satisfied by the wavelet-based approximation of $\hat{\mathbf{X}}$. Central Limit Theory can then be applied to derive the asymptotic normality of the wavelet-based functional estimator. Consequently, the following asymptotic functional confidence interval can be obtained, when K and N go to infinity:

$$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{W\hat{X}}_i - 3\sqrt{\mathrm{trace}(\widehat{R}_{\mathbf{W\hat{X}}})}, \frac{1}{N}\sum_{i=1}^{N}\mathbf{W\hat{X}}_i + 3\sqrt{\mathrm{trace}(\widehat{R}_{\mathbf{W\hat{X}}})}\right],$$

where

$$\widehat{R}_{\mathbf{W\hat{X}}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{W\hat{X}}_i(\cdot)\otimes\mathbf{W\hat{X}}_i(\star), \quad \mathbf{W\hat{X}}_i = \frac{1}{K}\sum_{k=1}^{K}\mathbf{W\hat{X}}_i^{(k)}, \quad i = 1,\ldots N,$$

and $\widehat{\mathbf{WX}}_i^{(k)}$ is defined from the wavelet-based orthogonal transform of the kth functional element $\mathbf{Y}_i^{(k)}$ of the sample $\mathbf{Y}_i^{(1)},\ldots,\mathbf{Y}_i^{(K)}$, of size K, available at time i, with $i = 1,\ldots,N$. Here, $\mathrm{trace}(\widehat{R}_{\mathbf{W\hat{X}}})$ denotes the trace of the empirical covariance operator $\widehat{R}_{\mathbf{W\hat{X}}}$.

4 Data analysis

In this section, we consider the filtering problem from fMRI data, in terms of the above-described wavelet-based orthogonal approximation of the corresponding functional least-squares estimator. A sequence of 9 images is considered. Estimated functional variances (EFV) of the wavelet-based functional spatial estimates are displayed in Table (1). Figure (1) shows an observed image and its functional estimation at the first time instant.

Acknowledgments. This work has been supported in part by projects MTM2009-13393 of the DGI, MEC, and P09-FQM-5052 of the Andalousian CICE, Spain.

Time	T=1	T=3	T=3	T=4	T=5	T=6	T=7	T=8	T=9
EFV	0.1306	0.1083	0.0259	0.0137	0.0146	0.0908	0.0386	0.1911	0.2032

Table 1: Estimated functional variances.

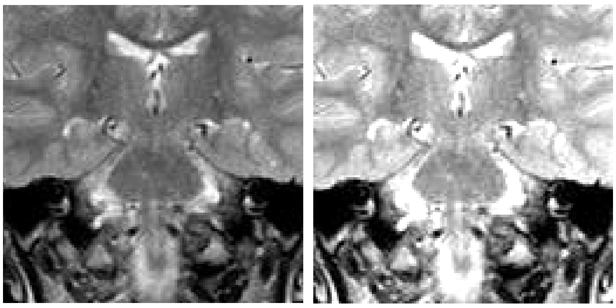

Figure 1: *Observed image at first time considered* $t = 1$ *(left), wavelet-based filtered image (right).*

References

[1] Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.J. (1995). Statistical parametric maps in fucntional imaging: A general approach. *Human Brain Mapping* **2**, 189–210.

[2] Friston, K.J. (2007). *Statistical parametric mapping: the analysis of functional brain images*. Academic Press, inc.

[3] Lazar, N.A. (2008). The statistical analysis of functional MRI data. *Statistics for biology and Health*. Springer.

[4] Lindquist, M. A. (2008). The Statistical Analysis of fMRI Data. *Statistical Science* **23,** 439–464.

[5] Ruiz-Medina, M.D. and Fernández-Pascual, R. (2010). Spatiotemporal filtering from fractal spatial functional data sequence. Stochastic Environmental Research and Risk Assessment DOI: 10.1007/s00477-009-0343-x

[6] Wager, T. D., Hernandez, L., Jonides, J., and Lindquist, M. (2007). Elements of functional neuroimaging. In J. T. Cacioppo, L. G. Tassinary and G. G. Berntson (Eds.), *Handbook of Psychophysiology 4th ed.*, 19–55. Cambridge University Press.

[7] Woolrich, M.W., Beckmann, C.F., Nichols, T. E. and Smith, S.M. (2009). Statistical Analysis of fMRI Data. *fMRI Techniques and Protocols* **41**, 179–236.