Modelling the fMRI response using smooth FIR filters

Methods # 210

Finn Årup Nielsen[1], Cyril Goutte[2] and Lars Kai Hansen[1]

[1]Informatics and Mathematical Modelling, Technical University of Denmark

International Neuroimaging Consortium

[2]INRIA Rhône-Alpes, France

Introduction

We present a model — the *smooth FIR filter* — suitable for flexible modelling of smooth linear time-invariant transfer functions. We apply the method for modelling the fMRI response.

Background

For linear modelling of the fMRI response the general linear model can be considered [5]:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{U}, \tag{1}$$

where $\mathbf{Y}(N \times Q)$ represents N fMRI scans each with Q voxels, $\mathbf{X}(N \times P)$ is a design matrix with P design variables, and $\mathbf{B}(P \times Q)$ is a matrix of parameters. Working independently with each voxel:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{u}, \tag{2}$$

the ordinary least squares (OLS) estimate is

$$\hat{\mathbf{b}}_{\mathsf{OLS}} = \left[\mathbf{X}^\mathsf{T}\mathbf{X}\right]^{-1}\mathbf{X}^\mathsf{T}\mathbf{y}. \tag{3}$$

When the fMRI response is to be fitted (some of) the design variables should represent a flexible set of basis functions, e.g., Dirac delta functions — the so-called FIR filter [7], sinusoids [2], or a mixture of fixed gamma densities with derivatives [1]. If the model should be flexible it is necessary to use many basis functions. Too many basis functions will on the other hand *overfit*, there is a *large variance of the parameter estimate* and in classical statistical test *power is lost*.

Regularised estimation

Overfitting is particularly a problem in connection with the FIR filter as it requires one design variable for each scan. Considering a filter length of 30 seconds with a TR = 333ms (a single slice) the resulting filter will require 90 design variables (filter coefficients) to represent the fMRI response and the filter will typically overfit. With *regularisation* overfitting can be reduced while a large filter length can be maintained. Principal component regression (PCR) regularises by confining the estimation space to a subspace of $\mathbf{X}$ [6, 8]:

$$\hat{\mathbf{b}}_{\mathsf{PCR}} = \mathbf{VL}^{-2}\mathbf{V}^\mathsf{T}\mathbf{X}^\mathsf{T}\mathbf{y} \tag{4}$$

$$\mathbf{ULV}^\mathsf{T} = \mathsf{svd}_K\left[\mathbf{X}\right]. \tag{5}$$

Ridge regression (RR) regularises by restricting the size of the parameters [4]:

$$\hat{\mathbf{b}}_{\mathsf{RR}} = \left[\mathbf{X}^\mathsf{T}\mathbf{X} + \lambda\mathbf{I}\right]^{-1}\mathbf{X}^\mathsf{T}\mathbf{y}. \tag{6}$$

These are not optimal for fMRI and here we propose *Smooth FIR* that regularises by penalising large differences between neighbouring filter coefficients b_i and b_j:

$$\hat{\mathbf{b}}_{\mathsf{SFIR}} = \left[\mathbf{X}^\mathsf{T}\mathbf{X} + \lambda\mathbf{R}\right]^{-1}\mathbf{X}^\mathsf{T}\mathbf{y} \tag{7}$$

$$\mathbf{R} = \mathbf{\Omega}^{-1}, \quad \Omega_{ij} = \exp\left[-\frac{1}{2l^2}(i-j)^2\right], \tag{8}$$

where λ is the amount of regularisation and l is the characteristic length. With $l \rightarrow 0$ ridge regression is recovered.

Probabilistic formulation

The estimation procedure can also be developed from probabilistic assumptions: Assuming the noise is isotropic and independent Gaussian distributed — $u \sim \mathcal{N}(0, \sigma_u^2)$ — the likelihood is:

$$p(\mathbf{y}|\mathbf{X},\mathbf{b},\sigma_u^2) = (2\pi\sigma_u^2)^{-\frac{N}{2}} \exp\left[-\frac{(\mathbf{Xb}-\mathbf{y})^\mathsf{T}(\mathbf{Xb}-\mathbf{y})}{2\sigma_u^2}\right]. \tag{9}$$

A multivariate Gaussian prior on the P parameters is specified

$$p(\mathbf{b}|\mathbf{\Sigma_{bb}}) = |2\pi\mathbf{\Sigma_{bb}}|^{-\frac{P}{2}} \exp\left[-\frac{\mathbf{b}^\mathsf{T}\mathbf{\Sigma}_{\mathbf{bb}}^{-1}\mathbf{b}}{2}\right] \tag{10}$$

The posterior is found through Bayes formula multiplying the likelihood and prior:

$$p(\mathbf{b}|\mathbf{X},\mathbf{y},\mathbf{\Sigma_{bb}},\sigma_\mathbf{u}^2) \propto \tag{11}$$

$$p(\mathbf{y}|\mathbf{X},\mathbf{b},\sigma_\mathbf{u}^2) \times p(\mathbf{b}|\mathbf{\Sigma_{bb}}) \propto \tag{12}$$

$$\exp\frac{(\mathbf{b}-\hat{\mathbf{b}})^\mathsf{T}\left(\mathbf{X}^\mathsf{T}\mathbf{X}+\sigma^2\mathbf{\Sigma}_{\mathbf{bb}}^{-1}\right)(\mathbf{b}-\hat{\mathbf{b}})}{-2\sigma^2}, \tag{13}$$

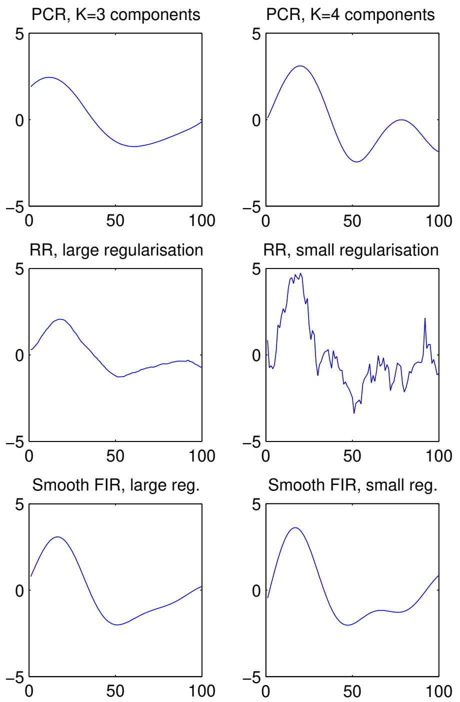
where $\hat{\mathbf{b}}$ is the maximum a posteriori (MAP) estimate:

$$\hat{\mathbf{b}} = \left[\mathbf{X}^\mathsf{T}\mathbf{X} + \sigma_u^2\mathbf{\Sigma}_{\mathbf{bb}}^{-1}\right]^{-1}\mathbf{X}^\mathsf{T}\mathbf{y} \tag{14}$$

By comparison with equations 7 and 8 the hyperparameters can be set to $\lambda = \sigma_\mathbf{u}^2/\nu$ and $\mathbf{\Omega} = \mathbf{\Sigma_{bb}}/\nu$, where ν controls the amount of regularisation compared to the noise level.

Example

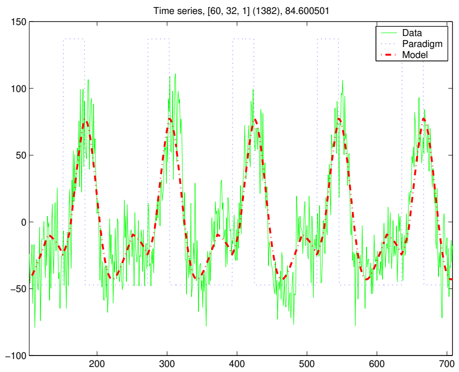
Egill Rostrup, Hvidovre hospital, Single slice fMRI 2D gradient echo EPI, B/FA/TE/TR = 1.5T/50°/66ms/333ms. The paradigm is full-field 8Hz-reversing checkerboard, 10 runs with 121 scans each.
Below is shown the the filter $\hat{\mathbf{b}}$ as modelled with PCR, ridge regression (RR) and smooth FIR on the same activated voxel. The filter length is 100 corresponding to 33 seconds.



... Example

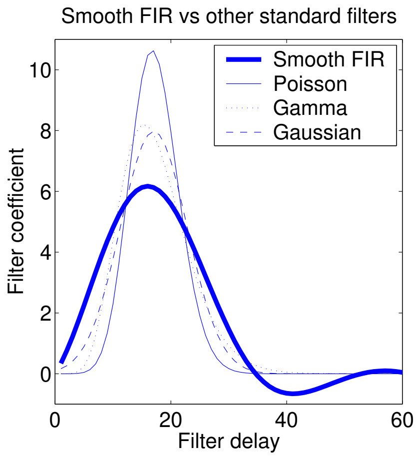
The estimation with PCR often results in filters with oscillations if too many components are used.

The ridge regression filter produces a noisy filter with small regularisation, and with large regularisation the filter parameters *shrink* towards zero: the amplitude of the filter will be smaller than optimal. This is also apparent in the figure above where the predicted fMRI signal ($\mathbf{X}\hat{\mathbf{b}} = \hat{\mathbf{y}}$, red curve) is smaller than the original signal ($\mathbf{y}$ green curve).



The smooth FIR (above) does not skrink the parameters significantly and does not oscillate as much as the PCR model.

... Example



If the smooth filter is compared to simple parametric fMRI response models, such as Poisson, gamma density and Gaussian, its flexibility is shown by its ability to model a post-activation undershoot — a characteristic also present in the default SPM99 fMRI response model [1].

Conclusion

The smooth FIR filter is a flexible linear model that can model chararistics of the fMRI response not found in simple parametric models, while it avoids the unstable prediction found in the ordinary FIR filter when large filters are necessary. The type of regularisation appears more appropriate to capture the form of the fMRI response than other types of regularisations.

More information

Part of the techniques are implemented in the *Lyngby Matlab toolbox* available at `http://hendrix.imm.dtu.dk/software/lyngby/` The method is described in full detail in [3].

Remarks, further work

- The hyperparameters λ (regularisation strength) and l (characteristic length) need to be determined depending on the amount of noise in the data and how much data there is. As a heuristics l can be set to 5-10 seconds.
- Classical statistical tests are not directly valid. An approximate test using the *effective* degrees of freedom might be developed.
- The "delay" of the hemodynamic response can be evaluated as the *group delay* of the filter:
$$\tau = \frac{\sum_i b_i \times i}{\sum_i i} \qquad (15)$$
- The method can also be used in connection with ANCOVA modelling where covariates with no relation to the fMRI response are present among the design varibles. In such a case the regularisation should not include those variables:
$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad (16)$$
- Other uses of the smooth FIR model are possible, e.g., in connection with perfusion MRI instead of principal component regression [8], and perhaps also for non-parametric modelling in connection with dynamical PET.

Acknowledgement

Egill Rostrup, Hvidovre Hospital provided the data. This project has been funded by The Danish Research Councils through THOR Center for Neuroinformatics and the Human Brain Project P20 MH57180 "Spatial and Temporal Patterns in Functional Neuroimaging". EU BIOMEDII

References

[1] Karl J. Friston. spm_hrf.m. Included in the SPM distribution, May 1999. Version 2.7.

[2] Karl J. Friston, Chris D. Frith, Richard S. J. Frackowiak, and Robert Turner. Characterizing dynamic brain responses with fMRI: A multivariate approach. *NeuroImage*, 2(2):166–172, June 1995.

[3] Cyril Goutte, Finn Årup Nielsen, and Lars Kai Hansen. Modelling the haemodynamic response in fMRI using smooth FIR filters. *IEEE Transactions on Medical Imaging*, 19(12):1188–1201, December 2000.

[4] Arthur E. Hoerl and Robert W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, February 1970.

[5] Kantilal Varichand Mardia, John T. Kent, and John M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979.

[6] W. F. Massy. Principal component analysis in exploratory data research. *Journal of the American Statistical Association*, 60:234–256, 1965.

[7] Finn Årup Nielsen, Lars Kai Hansen, Peter Toft, Cyril Goutte, Niels Jacob Sand Mørch, Claus Svarer, Robert Savoy, Bruce R. Rosen, Egill Rostrup, and Peter Born. Comparison of two convolution models for fMRI time series. *NeuroImage*, 5(4):S473, 1997.

[8] L. Østergaard, Robert M. Weiskoff, D. A. Chesler, C. Gyldensted, and Bruce R. Rosen. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. part I: Mathematical approach and statistical analysis. *Magnetic Resonance in Medicine*, 36(5):715–725, November 1996.