ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

NX-421 MINIPROJECT 1
VARIANT 3

GROUP G

Richard Gao
Giacomo Massini
Matteo Righi
Isabella Schøning

24th October 2022

1 INTRODUCTION

Functional magnetic resonance imaging (fMRI) is used to study neural activity to stimuli by measuring vascular response in the brain. Such measurements can help to explain underlying neural activity and can inform about the parts of the brain involved in processing certain tasks or stimuli. In order to analyze fMRI data, it is important to be able to accurately model the evoked hemodynamic response to a neural event (Lindquist et al. 2009).

The purpose of this project is to specify a fake sequence of stimuli events, model the blood oxygenation level-dependent (BOLD) signal for a single subject using the hemodynamic response function (HRF), and recover the true signal given the simulated time series for a single voxel using the general linear model (GLM).

2 METHODS

The context is to determine which areas of the brain are involved in processing faces. An experiment is first simulated in which a subject is shown alternating images of faces (visual stimuli) and of objects (control stimuli). The neural activation to the presentation of faces is modelled as a sequence of square waves: a high-amplitude square wave represents the activation corresponding to the faces stimulus and a lower-amplitude square wave represents the activation corresponding to the control stimulus. When no stimuli are present, the activation is zero. The parameters chosen for this experiment can be seen in the waveform in figure 1b: faces amplitude of 1, control amplitude of 0.1, faces and controls shown for 2 seconds, total trial duration is 10 seconds, total number of trials is 10.

We model the brain as an LTI system which takes as input the stimulus and outputs a corresponding hemodynamic response. The LTI system is characterized by its impulse response: the hemodynamic response function (HRF). In this case, the canonical HRF is formed by a difference of Gamma functions:

$$h(t) = A\left(\frac{t^{\alpha_1 - 1}\beta_1^{\alpha_1} e^{\beta_1 t}}{\Gamma(\alpha_1)} - c\frac{t^{\alpha_2 - 2}\beta_2^{\alpha_2} e^{\beta_2 t}}{\Gamma(\alpha_2)}\right) \tag{1}$$

where $\alpha_1 = 6$, $\alpha_2 = 16$, $\beta_1 = \beta_2 = 1$, $c = \frac{1}{6}$, and $A = 1$ is the amplitude of the HRF, recovered through model fitting. The HRF function is shown in figure 1a.

The hemodynamic response to our experiment is simulated by convolving the activation signal defined above with the HRF. The convolved signal corresponds to a simulation of the blood-oxygen level dependent (BOLD) signal, seen in figure 1c.

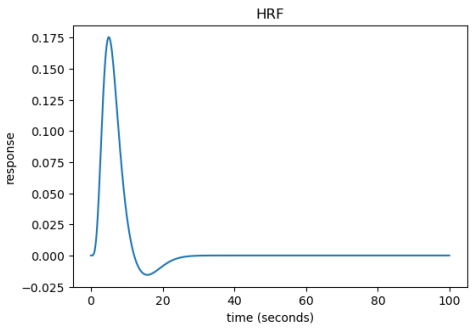
Next, noise is added to the BOLD signal to better simulate reality: in a real signal, there is noise coming from low-frequency drift, auto-correlation, or possibly some aliased physiological artifacts. In this simulation, Gaussian noise is added to the simulated BOLD signal with a standard deviation $\sigma = 0.15$. Note that by linearity of the convolution operator, it is equivalent to add noise before or after the convolution. This signal is denoted Y and can be seen in figure 1c.

3 RESULTS

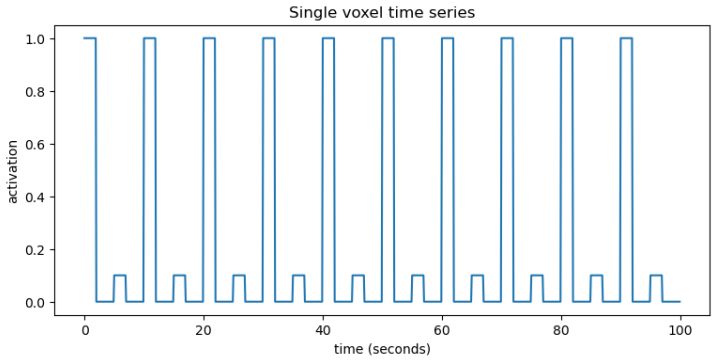
To recover the original signal from the noisy signal, a general linear model (GLM) is used, of the form:

$$Y = X\beta + \epsilon \tag{2}$$

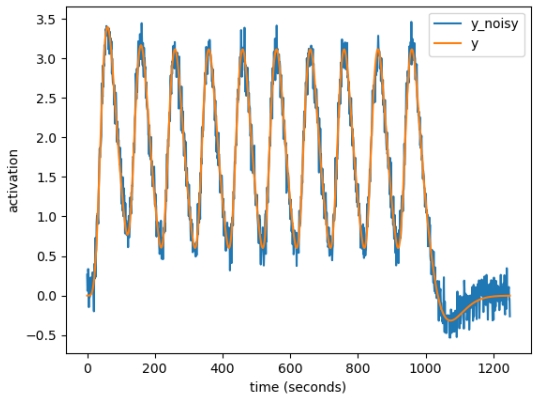
where Y is the simulated noisy BOLD voxel time series, X is the model or design matrix, ϵ is independent Gaussian noise, and β is a vector of parameters which are estimated to scale the model.



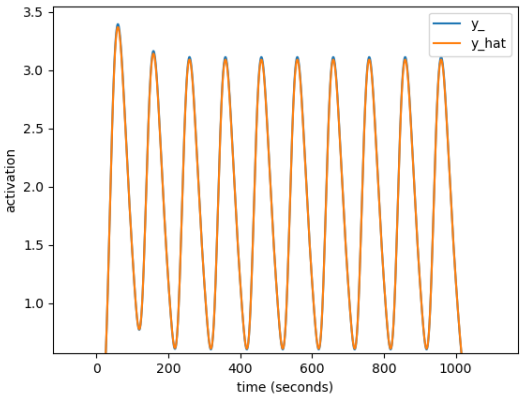
(A) Simulated HRF.



(B) Simulation of the time series for a single voxel.



(C) Simulated real and noisy BOLD signals.



(D) Reconstructed signal compared to real signal.

The design matrix has three columns and as many rows as there are time points in the voxel time series. The first column corresponds to the face stimuli: it is built by convolving the HRF with a square wave with amplitude 1 whenever a face is presented. The second column is constructed in analogous fashion, but with a square wave which is high in the presence of the control stimulus. The third column is a constant, i.e. a column of ones. For more detail, a graphical representation of the design matrix is shown in the Appendix in figure 2.

The optimal β vector is calculated using the Ordinary Least Squares (OLS) estimator:

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

which gives the estimator $\hat{\beta} = [0.98794,\ 0.098589,\ 0.01524]$. Each element is considered equivalent to a p-value, which gives, respectively, 99% and 9.9% probability that the first two regressors are generating the BOLD signal, and 1.5% chance that the third is.

The best estimate of the original BOLD signal can then be reconstructed as $\hat{y} = X\boldsymbol{\beta}$. The reconstructed signal is shown graphically in figure 1d. The residual error between the real and the reconstructed signals is a metric of how accurate the signal reconstruction has been. An L1-type metric is taken: $res = \sum_{t=0}^{T} |y_t - \hat{y}_t|$. This gives a residual error of 14.577.

4 DISCUSSION

The parameters of β can tell us whether a region is responsible for processing faces. Since the first regressor associated with faces is 99%, we can conclude that this voxel is indeed responsible for processing faces. Because the p-value for the third regressor is below the threshold 5%, we consider the column of constants statistically insignificant in generating the Y signal. This result is coherent with the construction of the experiment: indeed, the BOLD signal mainly consists of the faces stimulation signal, with a small contribution of the control stimulation signal. If the amplitude of control stimuli was set lower, the second regressor would be less significant in generating Y, as well.

We consider the effect of the α parameters in the HRF (equation 1). It can be shown that in the Gamma PDF, the α parameter affects the variance of the distribution: $\alpha \to 1^+$ gives a tightly peaked distribution, with a peak occurring close to time 0, and $\alpha \to \infty$ gives a shorter and wider distribution with a peak occurring later. Recall that the HRF consists of a difference of Gamma distributions; thus, α_1 affects the initial, larger peak corresponding to the initial hemodynamic overcompensation, and α_2 affects the subsequent, smaller dip corresponding to the hemodynamic negative overshoot. This is shown graphically in the Appendix in figure 3.

The effect of varying the signal amplitude is also considered, i.e. varying the A parameter of the HRF. We repeated the entire procedure explained in the previous sections with different values of A. Note that in the Python code, the random number generator seed was fixed so that noise would be consistent across runs. We found that the residual error was the same for different signal amplitude values. This follows from the fact that the error is related to the difference between the real and noisy signals, therefore it is not linked to the original signal amplitude. For certain applications (for example when using amplifiers), a more relevant metric might be residual error relative to the maximal signal amplitude. In this case, for larger signal amplitudes the relative error is less great.

We also consider the effect of varying the level of added noise. This was done as described in the previous paragraph, this time varying the standard deviation σ of the added Gaussian noise. We found that as σ is increased, the residual error increases linearly. This is logical, as, again, the residual error is related to the difference between the real and noisy signals.

Lastly, we consider the effect of varying the number of trials. As is the case in machine learning, with too few trials the model will be less accurate, but with too many trials the model starts to overfit. This is of less concern in our case because the data is regular and is disrupted only by the noise. Moreover, we are not implementing a testing set, so it is not possible to test for overfitting. To account for the fact that longer signals (i.e. with more trials) will necessarily have a higher summed error, we compare the residual error relative to the number of trials $\frac{res}{n.\ trials}$. It was found that as the number of trials is increased, the relative error decreases in non-linear fashion, which is consistent with our intuition.

Quantitative results are summarized in table 1 below:

A	$\frac{res}{A}$	σ	res	n. trials	$\frac{res}{n.\ trials}$
0.1	145.77	0.1	0.9718	1	6.8333
1	14.577	1	9.718	10	1.4577
10	1.4577	10	97.18	100	0.27643

TABLE 1
Performance varying different simulation parameters.

REFERENCES

Lindquist, Martin A. et al. (2009). 'Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling.' In: *NeuroImage* 45 (1 Suppl). ISSN: 10959572. DOI: 10.1016/J.NEUROIMAGE.2008.10.065.

5 APPENDIX

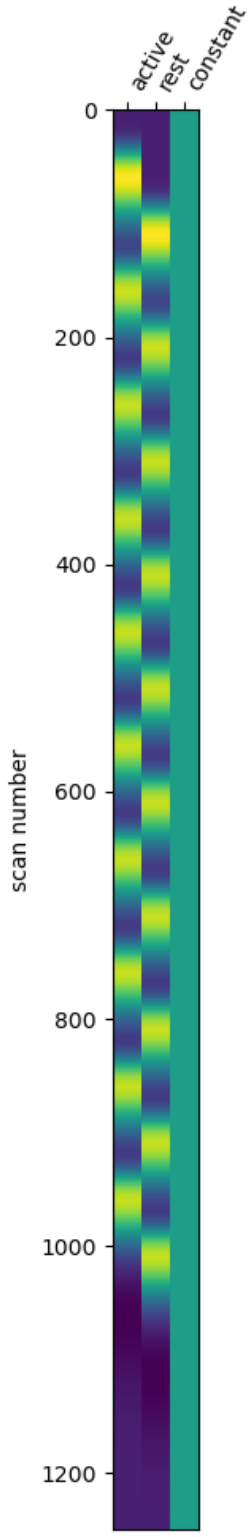


FIGURE 2
A graphical representation of the design matrix X where 'active' corresponds to an image of a face being shown and 'rest' corresponds to a control image.

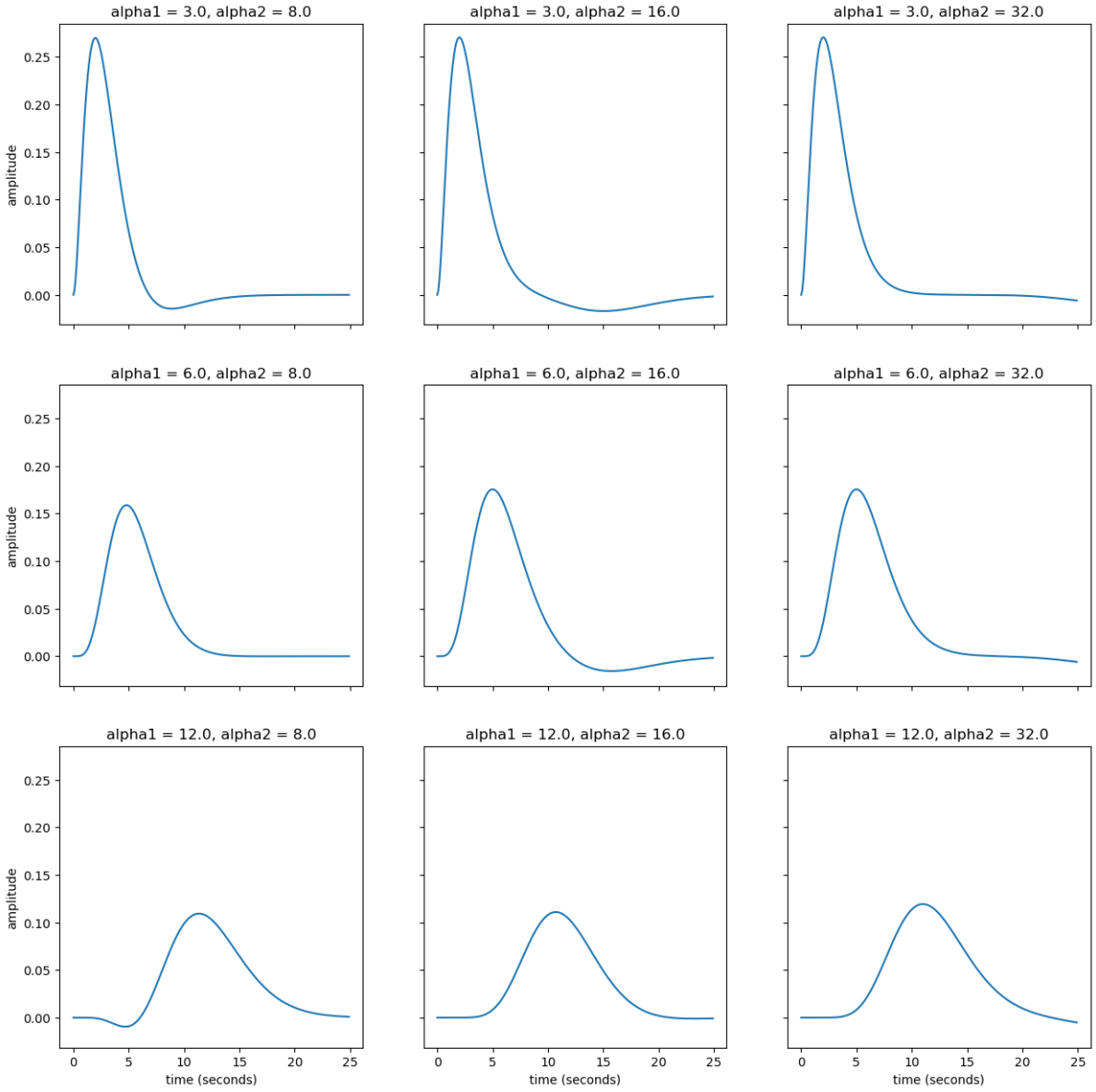


FIGURE 3
Effects of changing the HRF function parameters by $\alpha_1 = [0.5, 1, 2] * \alpha_1$ and $\alpha_2 = [0.5, 1, 2] * \alpha_2$.