Harvey et al., 2013, but with equations

Joram Soch

soch@cbs.mpg.de

Abstract

We review the mathematical formulation of the statistical model used in "Topographic Representation of Numerosity in the Human Parietal Cortex" (Harvey et al., *Science*, 2013; vol. 341, pp. 1123-1126; DOI: 10.1126/science.1239052) and provide actual equations for the entire model, with notation that is easy to follow and common in cognitive neuroscience. Since this study does not report any equations, we occassionally reference "Population receptive field estimates in human visual cortex" (Dumoulin & Wandell, *NeuroImage*, 2008; vol. 39, pp. 647-660; DOI: 10.1016/j.neuroimage.2007.09.034) which describes the modelling framework on which the analysis was based.

Contents

1 Introduction

Harvey and colleagues presented a computational model for numerosity processing in human cerebral cortex (Harvey et al., 2013), measured with functional magnetic resonance imaging (fMRI) and analyzed using a population receptive field (pRF) model for visual perception by Dumoulin and Wandell (Dumoulin & Wandell, 2008). Briefly, measured fMRI signals were assumed to originate from neuronal activity following presented numerosity that follows log-Gaussian-shaped tuning functions, convolved with some hemodynamic response to yield a hemodynamic signal. This approach was later re-applied and extended by Harvey and colleagues (e.g. Harvey et al., 2011; Harvey et al., 2015; Harvey & Dumoulin, 2017; Paul et al., 2022).

Here, we review the statistical model formulation behind numerosity pRF models to facilitate simulation of fMRI signals compatible with specific tuning parameters in the context of our own research project on the EMergence of PRecISE numerosity representations in the human brain (EMPRISE).

2 Model specification

2.1 Notation

An fMRI numerosity processing experiment has been performed. Let $x_t \in \mathcal{X}$ be the presented numerosity at continuous time t. Here, $\mathcal{X} = \{1, 2, \ldots, L, M\}$ is the stimulus space where L is the number of low numerosities levels (Harvey et al., 2013: $L = 7$; EMPRISE: $L = 5$) and M is the high numerosity presented in intermissions (Harvey & EMPRISE: $M = 20$; Harvey & Dumoulin, 2017: $L \in \{10, 20\}$?[1]).

Furthermore, let y_{ji} be the measured BOLD signal in a single voxel during the i-th scan (EMPRISE: $i = 1, \ldots, 145$) belonging the j-run (EMPRISE: $j = 1, \ldots, 8$). We denote as $\bar{y}_j$ the signal which results from averaging across repeated "epochs" of identical stimulus presentation within each run (Harvey & EMPRISE: 4 epochs) and we denote as $\bar{y}$ the signal which results from averaging $\bar{y}_j$ across runs.

2.2 Neuronal model

We define a logarithmic Gaussian tuning function on stimulus x as

$$f_{\log}(x) = \exp\left[-\frac{1}{2}\frac{(\ln x - \mu)^2}{\sigma_x^2}\right] \tag{1}$$

and a linear Gaussian tuning function (cf. Dumoulin & Wandell, 2008, eq. 2[2]) as

$$f_{\lin}(x) = \exp\left[-\frac{1}{2}\frac{(x - \mu)^2}{\sigma_x^2}\right] \tag{2}$$

---

[1] Harvey & Dumoulin, 2017, p. 7: "The numerosities one to seven were first presented in ascending order, followed by 16.8 seconds showing 10 circles (sic!), followed (sic!) seven to one in descending order, followed by 16.8 seconds with 20 circles." However, this is incompatible with their Fig. 1 which shows a single high numerosity of 20.

[2] There is a typesetting error in this equation: The minus sign needs to be inside the parantheses.

where $\exp(\mu)$ (logarithmic) or μ (linear) corresponds to the preferred numerosity of the tuned population and $\exp(\sigma_x)$ (logarithmic)[3] or σ_x (linear) corresponds to the tuning width of the tuned population.

Instead of using the standard deviation, tuning functions are preferably parametrized in terms of the full width at half maximum (FWHM) that can be directly calculated from the standard deviation[4]:

$$\omega = \text{FWHM} = 2\sqrt{2}\ln 2\sigma \ . \tag{3}$$

Then, neuronal activity at time t is assumed to be (cf. Dumoulin & Wandell, 2008, eq. 3)

$$\begin{aligned} z_t &= \sum_{x \in \mathcal{X}} [x = x_t] \cdot f_{\log}(x) \\ &= f_{\log}(x_t) \end{aligned} \tag{4}$$

where the Iverson bracket notation $[x = x_t]$ indicates presence of stimulus x_t at time t.

2.3 Hemodynamic model

We define $h(t)$ as the hemodynamic response function (HRF) at post-stimulus time t, e.g. obtained from SPM using HRF parameters θ:

$$h(t) \leftarrow \texttt{spm\_hrf}(\theta) \ . \tag{5}$$

Then, the hemodynamic signal at time t is assumed to be a convolution of the neural activity with the HRF (cf. Dumoulin & Wandell, 2008, eq. 4):

$$\begin{aligned} s_t &= z_t * h(t) \\ &= \int_0^t z_\tau \cdot h(t - \tau)\,\mathrm{d}\tau \ . \end{aligned} \tag{6}$$

Finally, the measured hemodynamic signal is assumed to be a linear combination of the underlying hemodynamic signal and zero-mean normally distributed measurement noise (cf. Dumoulin & Wandell, 2008, eq. 1[5])

$$y = \beta_s\, s + \varepsilon, \ \varepsilon_i \sim \mathcal{N}(0, \sigma^2), \ i = 1, \dots, n \tag{7}$$

where y is an $n \times 1$ vector representing the measured BOLD signal (n = number of scans), s is an $n \times 1$ vector of the hemodynamic signal s_t, sampled at the acquisition times of y, β_s is a scaling factor and σ is the noise variance.

---

[3]This is not yet correct. Tuning width is transformed from logarithmic space to linear space in a different way than preferred numerosity. For an implementation, see here: https://github.com/SkeideLab/EMPRISE/blob/akieslinger/data/code/harveymodel.py#L418 (function "sigma2fhwm").

[4]For a proof of the relationship, see hee: https://statproofbook.github.io/P/norm-fwhm.

[5]To be precise, Dumoulin and Wandell make no assumptions about the noise distribution.

3 Model estimation

3.1 Estimation of tuning parameters

Let μ and ω be candidate tuning parameters (preferred numerosity and tuning width, i.e. FWHM) describing the potential neuronal response of a single voxel relative to presented stimuli x_t. We denote the resulting tuning function (see eq. 1) as $f_{\log}(x_t; \mu, \omega)$, the expected neuronal signal (see eq. 4) as $z_t(\mu, \omega)$ and the expected hemodynamic signal (cf. eq. 6) as $s(\mu, \omega)$.

Then, we can estimate the scaling factor by comparing the expected against the measured hemodynamic signal

$$\hat{\beta}_s(\mu, \omega) = \arg\min_{\beta_s} \sum_{i=1}^{n} [y_i - \beta_s s_i(\mu, \omega)]^2 \tag{8}$$

and calculate the residual sum of squares (RSS) as a function of the estimated scaling factor (cf. Dumoulin & Wandell, 2008, eq. 5):

$$\text{RSS}(\mu, \omega) = \sum_{i=1}^{n} \left[y_i - \hat{\beta}_s(\mu, \omega) s_i(\mu, \omega) \right]^2 . \tag{9}$$

The estimated preferred numerosity $\hat{\mu}$ and estimated tuning width $\hat{\omega}$ are then obtained as those candidate tuning parameters that minimize the RSS:

$$(\hat{\mu}, \hat{\omega}) = \arg\min_{(\mu, \omega)} \text{RSS}(\mu, \omega) . \tag{10}$$

3.2 Extensions by the authors

In order to account for different hemodynamic response functions between subjects, but not between regions or voxels, Harvey and colleagues report (Harvey et al., 2013, suppl. p. 9; Harvey & Dumoulin, 2017, p. 7) to have estimated subject-wise HRF parameters after an initial estimation of tuning parameters (which was based on default HRF parameters) and then re-estimating tuning parameters (using the updated HRF parameters). This estimation scheme can be summarized as follows[6]:

$$
\begin{aligned}
h_0(t) &\leftarrow \texttt{spm\_hrf}(\theta_0) \\
(\hat{\mu}_0, \hat{\omega}_0) &= \arg\min_{(\mu, \omega)} \text{RSS}(\mu, \omega; h_0) \\
\theta &\leftarrow \{y, x, \hat{\mu}_0, \hat{\omega}_0\} \\
h(t) &\leftarrow \texttt{spm\_hrf}(\theta) \\
(\hat{\mu}, \hat{\omega}) &= \arg\min_{(\mu, \omega)} \text{RSS}(\mu, \omega; h)
\end{aligned}
\tag{11}
$$

where θ_0 are the default HRF parameters and subscript 0 signifies initial estimates.

---

[6]It is obvious that the third line of these equations is dubious and needs clarification. For an implementation, see here: https://github.com/SkeideLab/EMPRISE/blob/akieslinger/data/code/harveymodel. py#L549 (function "estimate_hrf").

3.3 Possible further extensions

It is very unlikely that equation (7) is a realistic model of preprocessed, but otherwise unaltered fMRI data. In particular, this model does not account for a baseline level of hemodynamic activity and does allow experimental conditions of no interest and nuisance variables, possibly influencing the preprocessed signal. We assume that these factors can be summarized as a design matrix

$$X(\mu, \omega) = \begin{bmatrix} s(\mu, \omega) & X_c & 1_n \end{bmatrix} \tag{12}$$

where $s(\mu, \omega)$ is identical to the s in equation (7), X_c is an $n \times c$ matrix of confounds and 1_n is an $n \times 1$ vector of ones, such that the model from (7) changes to

$$y = X\beta + \varepsilon, \ \varepsilon \sim \mathcal{N}(0, \sigma I_n) \ . \tag{13}$$

Then, regression coefficients can be estimated using the ordinary least squares equation

$$\hat{\beta}(\mu, \omega) = (X^\mathrm{T} X)^{-1} X^\mathrm{T} y \tag{14}$$

and estimated tuning parameters would be given as minimizing the RSS:

$$(\hat{\mu}, \hat{\omega}) = \underset{(\mu, \omega)}{\arg\min} \left[y - X\hat{\beta}(\mu, \omega) \right]^\mathrm{T} \left[y - X\hat{\beta}(\mu, \omega) \right] \ . \tag{15}$$

Another restriction of (7) consists in the fact that it assumes independent and identically distributed (i.i.d.) error terms ε_i and thus does not allow for error covariance which typically exists in the form of serial correlations in fMRI. We assume that such serial correlations are given in the form of a temporal covariance matrix V, e.g. obtained using SPM's restricted maximum likelihood (ReML) approach:

$$V \leftarrow \texttt{spm\_reml}(Y, X) \tag{16}$$

where Y is an $n \times v$ matrix of fMRI signals across all voxels inside a region of interest (ROI) or in the whole brain, such that the model from (7) changes to

$$y = X\beta + \varepsilon, \ \varepsilon \sim \mathcal{N}(0, \sigma V) \ . \tag{17}$$

Then, regression coefficients can be estimated using the weighted least squares equation

$$\hat{\beta}(\mu, \omega) = (X^\mathrm{T} V^{-1} X)^{-1} X^\mathrm{T} V^{-1} y \tag{18}$$

and estimated tuning parameters would be given as minimizing the weighted RSS:

$$(\hat{\mu}, \hat{\omega}) = \underset{(\mu, \omega)}{\arg\min} \left[y - X\hat{\beta}(\mu, \omega) \right]^\mathrm{T} V^{-1} \left[y - X\hat{\beta}(\mu, \omega) \right] \ . \tag{19}$$