Signal Processing and Statistical Analysis for Event-Related fMRI

Prepared by: Douglas N. Greve

February 22, 1999

1 Introduction

In an fMRI experiment, one or more different types of stimuli are presented to the subject and the hemodynamic response (HDR), is measured for a given time after presentation. In a *block* experimental design, the time between the offset of one stimulus and the onset of the next is longer than it takes for the hemodynamic response to regain equilibrium. In an *event-related* design, one stimulus can appear before the hemodynamic response from the previous stimulus has decayed away. This allows for more interesting psychological experiments but creates overlap in the responses. This overlap can confound the results. This document describes the signal processing and statistical analysis of the fMRI signal for event-related designs. This processing stream accounts for the overlap in the hemodynamic response as well as compensates for the temporal correlation in the fMRI noise.

2 Model of the Hemodynamic Response

The hemodynamic response **at a single voxel** to the same stimulus (or same type of stimulus) presented several times over an experiment is assumed to be governed by following linear time-invariant system:

$$y(t) = x(t) * h(t) + n(t) \tag{1}$$

where $y(t)$ is the fMRI signal (ie, the observable), $x(t)$ is a pulse train which has a value of 1 at the onset time of each presentation and zero everywhere else[1], $h(t)$ is the (unknown) hemodynamic response to a single presentation of the stimulus, $n(t)$ is additive noise distributed $N(0, \Sigma_n)$, and $*$ is the convolution operator. Furthermore, we assume the noise covariance matrix Σ_n can been factored into two components:

$$\Sigma_n = \sigma_n^2 C_n, \tag{2}$$

where σ_n^2 is the (scalar) variance of the noise at the voxel and C_n is a symmetric, positive-definite matrix with ones on the diagonal (the "normalized" covariance matrix). It is further assumed that, while each voxel will have its own σ_n, C_n will be the same across all voxels (or at least all voxels within the skull). If the noise is white, then $\Sigma_n = \sigma_n^2 I$, where I is the identity matrix. When the stimulus is brief, $h(t)$ becomes the hemodynamic impulse response.

[1] The list of the type of stimulus presented at a certain time is the content of the "parfile".

If the stimulus sequence has different types of stimuli (or conditions) that generate a different HDR, it is assumed that the HDRs combine linearly according to the extension of equation **??**:

$$y(t) = x_1(t) * h_1(t) + x_2(t) * h_2(t) + \ldots + x_{N_c}(t) * h_{N_c}(t) + n(t) \tag{3}$$

where N_c is the number of different types of stimuli, $x_i(t)$ is the presentation sequence for stimulus type i, and $h_i(t)$ is the hemodynamic response to a stimulus type i.

The continuous-time equation above can be converted into the following discrete-time, matrix model of the fMRI signal:

$$y(k) = X_1 h_1(k) + X_2 h_2(k) + \ldots + X_{N_c} h_{N_c}(k) + n(k), \tag{4}$$

where k ranges from 0 to N_{tp} (the number of time-points or observations), X_i is the stimulus convolution matrix (SCM) for stimulus type i, and $h_i(k)$ is the continuous-time HDR sampled at a delay $TR * k$ after the onset of a stimulus of type i, where TR is the time-between observations. The HDR is sampled over a time window of T_{HDR} resulting in $N_h = T_{HDR}/TR$ samples. This is a moving-average (MA) or finite-impulse response (FIR) model.

X_i, the SCM, is an asymmetric toeplitz matrix of ones and zeros. It has dimension $N_{tp} \times N_h$ (the number of time-points by the number of delays) and is constructed from the stimulus sequence $x_i(t)$. Let $s_i(k)$ be the stimulus sequence sampled at the TR. $s_i(k)$ is then a vector of ones and zeros. The columns of X_i are shifted version so of $s_i(k)$. Thus, the first column of x_i is identical to $s_i(k)$, where k becomes the row number. The second column has zero on the first row; the remainder of the rows are assigned the values $s_i(k), 1 \leq k \leq N_t - 1$. The third column has zeros on the first two rows, etc.

Equation **??** can be further consolidated in matrix notation:

$$y = Xh + n, \tag{5}$$

where $X = [X_1 X_2 \ldots X_{N_c}]$, ie a horizontal concatenation of the SCMs of the individual conditions, and h is the vertical concatenation of the individual HDRs. X has dimension $N_{tp} \times N_{ch}$ where $N_{ch} = N_c N_h$ is the total number of hemodynamic coefficients (ie, across all conditions and delays).

3 HDR Estimation

In Equation **??**, we know X and y and seek to estimate h. The hemodynamic response is obtained by solving

$$\min \parallel B_n^{-1}(Xh - y) \parallel_2 \tag{6}$$

where $B_n = C_n^{\frac{1}{2}}$ is the Cholesky factorization of C_n. This is the *generalized* least squares problem, which has the solution

$$h^* = (X^T C_n^{-1} X)^{-1} X^T C_n^{-1} y. \tag{7}$$

The *estimation error* is the difference between the estimated HDR and the actual HDR:

$$e_{h^*} = h - h^* = (X^T C_n^{-1} X)^{-1} X^T C_n^{-1} n \tag{8}$$

This should not be confused with the *residual error* which is

$$e_{y^*} = y - y^* = y - Xh^* \tag{9}$$

The generalized least squares solution minimizes the estimation error whereas the classical least squares solution minimizes the residual error. In the special case where the noise is white (ie, $C_n = I$), the same h^* minimizes both.

h^* is a random variable and will have a covariance matrix:

$$\Sigma_{h^*} = E(e_{h^*} e_{h^*}^T) = \sigma_n^2 (X^T C_n^{-1} X)^{-1} \tag{10}$$

Note that the RMS estimation error is equal to $trace(\Sigma_{h^*})$ which is independent of the actual value of h.

4 Noise Covariance Matrix Estimation

Estimation of the noise covariance matrix Σ_n requires two levels of processing. First we assume that the noise is white, $\Sigma_n = \sigma_n^2 I$, in which case our estimate from equation **??** becomes:

$$\bar{h} = (X^T X)^{-1} X^T y, \tag{11}$$

and the residuals are

$$e_{\bar{y}} = y - \bar{y} = y - X\bar{h} = (I - X(X^T X)^{-1} X) n. \tag{12}$$

For well counter-balanced sequences and $N_{tp} >> N_{ch}$, the elements of the matrix $X(X^T X)^{-1} X$ are small, and so, to a first approximation, the the residuals are a good estimate of the noise and so we will use the residuals to estimate C_n. Since C_n is the normalized covariance matrix for all voxels within the skull, we will compute it in the following way. First, the within skull voxels are identified as all the voxels greater than mean voxel value computed over space and time for a given slice over a given run. For each within-skull voxel, the normalized, unbiased autocorrelation of the residuals is computed. Normalizing forces the zero-delay coefficient to be unity. The global correlation function is computed by averaging each delay coefficient across all within-skull voxels. This global autocorrelation function is then used to fit the parameters in the following model:

$$\bar{R}_e(k) = \begin{cases} 1 & k = 0 \\ (1-\alpha)\rho^k & 0 < |k| \leq k_{max} \\ 0 & |k| > k_{max} \end{cases} \tag{13}$$

where k_{max} is used to limit the delay range over which the data are fit. For TR of 2 seconds, k_{max} is typically set to 10. Once α and ρ have been computed, a "synthetic" autocorrelation function is computed from equation **??**. $\bar{C}_n$ is then generated as a Toeplitz matrix using the synthetic function as a seed vector.

This is somewhat of a convoluted process, however, it is motivated by several factors:

- Noise from biological material has different spectral properties than that from air.
- We assume that C_n is *not* a random variable and so it should not depend upon how it is measured. We approximate this by fitting only two parameters to an average computed over a large number of voxels so that measurement artifacts will have little effect at any particular voxel.

We substitute this estimate of the noise covariance matrix into the formulas for the average and covariance of the HDR:

$$\hat{h} = (X^T \bar{C}_n^{-1} X)^{-1} X^T \bar{C}_n^{-1} y. \tag{14}$$

The estimation error is

$$e_{\hat{h}} = h - \hat{h} = (X^T \bar{C}_n^{-1} X)^{-1} X^T \bar{C}_n^{-1} n \tag{15}$$

$$\Sigma_{\hat{h}} = \hat{\sigma}_n^2 (X^T \bar{C}_n^{-1} X)^{-1}, \tag{16}$$

where $\hat{\sigma}_n^2$ is the variance of $e_{\hat{h}}$. Note that $\Sigma_{\hat{h}}$ can be factored into a voxel-dependent scalar and a voxel-independent matrix. This is important for practical reasons has a matrix does not need to be stored for every voxel.

5 Detrending

It is common for the fMRI signal offset to drift linearly and/or quadratically with time for non-physiological reasons. As this drift can skew the results, we need a way to remove both the offset and the offset drift. We refer to the remove of the offset and drift as *detrending*. To implement the detrending process, we first modify the model of the hemodynamics (equation **??**) to include offset and drift:

$$y(t) = x(t) * h_(t) + n(t) + a_0 + a_1 t + a_2 t^2 + \ldots + a_{N_{dt}-1} t^{N_d t - 1} + n(t), \tag{17}$$

where the value a_i is the scalar coefficient of the i^{th} trend. In the above equation, trends up to order N_{dt} are represented, where the first trend is a simple offset.

In matrix notation, equation **??** can be represented by:

$$y_{pre} = X_{dt} h_{dt} + n, \tag{18}$$

where y_{pre} is used to indicate the fMRI signal prior to detrending,

$$X_{dt} = \begin{bmatrix} X & t_s^0 & t^1 & \dots & t_s^{N_{dt}} \end{bmatrix} = [XD] \tag{19}$$

and

$$h_{dt} = \begin{bmatrix} h & a_0 & a_1 & \dots & a_{N_{dt}} \end{bmatrix} = \begin{bmatrix} h & a \end{bmatrix} \tag{20}$$

and t_s is a column vector of sample times ,ie, $t_s = \begin{bmatrix} 0 & TR & 2TR & \dots & (N_{tp}-1)TR \end{bmatrix}$. The hemodynamic responses are fit simultaneously with the trends (assuming white noise).

$$\tilde{h}_{dt} = (X_{dt}^T X_{dt})^{-1} X_{dt}^T y_{pre} \tag{21}$$

The trend can now be removed by:

$$y = y_{pre} - D\tilde{h}_{dt} \tag{22}$$

y is now the de-trended fMRI signal, and processing can proceed as indicated above with the one caveat that the degrees of freedom must be reduced by N_{dt}.

6 Processing Stream Summary

For multiple runs within a single session, the analysis proceeds according to the following steps. Each slice is processed separately and independently.

1. Compute stimulus convolution matrix for each run, X_r.
2. Detrend each run (equation **??**).
3. Estimate the hemodynamic response assuming white noise over the entire session according to the formula:

$$\bar{h}_S = \left(\sum_{r=1}^{N_r} X_r^T X_r \right)^{-1} \left(\sum_{r=1}^{N_r} X_r^T y_r \right) \tag{23}$$

4. Compute the residual errors for each run

$$e_{\bar{y},r} = y_r - X_r \bar{h}_S \tag{24}$$

5. Compute the estimated noise covariance matricies, $\bar{C_{n,r}}$, for each run.
6. Recompute the session hemodynamic response estimate

$$\hat{h}_s = \left(\sum_{r=1}^{N_r} X_r^T \bar{C}_{n,r}^{-1} X_r \right)^{-1} \left(\sum_{r=1}^{N_r} X_r^T \bar{C}_{n,r}^{-1} y_r \right), \tag{25}$$

7. Recompute the residual error at each voxel:

$$e_{\hat{h}_r} = y_r - X_r \hat{h}_s \tag{26}$$

8. Compute the variance of the residual over the entire session for each voxel:

$$\hat{\sigma}^2_{n,s} = \frac{\sum_{r=1}^{N_r} e^T_{\hat{h}_r} \bar{C}^{-1}_{n,r} e_{\hat{h}_r}}{N_{dof}}, \tag{27}$$

where the degrees of freedom, $N_{dof} = N_r * N_{tp} - N_{ch}$.

9. Compute the session hemodynamic covariance matrix:

$$\Sigma_{\hat{h}_s} = \hat{\sigma}^2_{n,s} \left(\sum_{r=1}^{N_r} X^T_r \bar{C}^{-1}_{n,r} X_r \right)^{-1} = \sigma^2_{\hat{h}_s} C_{\hat{h}_s} \tag{28}$$

where $\Sigma_{\hat{h}_s}$ can be factored into a voxel-dependent scalar and a voxel-independent matrix.

7 Statistical Inference

The goal of the fMRI experiment is to determine whether a location in the brain is becoming active in response to a particular stimulus type or whether one type activates a particular region more than another. Tests can also be restricted by range of post-stimulus delay. For example, the null hypothesis could be that none of the conditions at any delay generate a hemodynamic response that is significantly different than zero:

$$H_0 : \|\hat{h}\| = 0 \tag{29}$$

This can be tested using an F-test:

$$F(N_{ch}, N_{dof}) = \frac{\hat{h}^T C^{-1}_{\hat{h}} \hat{h}}{N_{ch} \sigma^2_{\hat{h}}} \tag{30}$$

In general, one may want to restrict the test to a subset of delays and conditions or combinations of conditions. For example, $H_0 : \|\hat{h}_1 - \hat{h}_2\| = 0$ or $H_0 : \|\hat{h}_3(3:6)\| = 0$. Mathematically, this is equivalent to multiplying $\hat{h}$ by a matrix R and testing the norm of the result.

$$q = R\hat{h}, H_0 : \|q\| = 0 \tag{31}$$

with the corresponding F test

$$F_q(N_q, N_{dof}) = \frac{q^T (R C_{\hat{h}} R^T)^{-1} q}{N_q \sigma^2_{\hat{h}}}, \tag{32}$$

where N_q is the number of rows in R. The restriction matrix, R, is typically a matrix of ± 1's and 0's. The number of columns of R must be equal to N_{ch}. In the example given in equations **??** and **??**, R would simply be an identity matrix of dimension N_{ch}. When R is a vector, then the F test of equation **??** reduces to a t test. Note that equation **??** can again be factored into a voxel-dependent scalar and a voxel-independent matrix.

8 Table of Variable Names

y	fMRI signal for a single voxel
y_{pre}	fMRI signal for a single voxel before detrending
X	Stimulus convolution matrix (SCM)
D	Trend Matrix
X_{dt}	Stimulus convolution matrix with trend matrix included
h	"True" HDR.
h_{dt}	"True" HDR with trend coefficients
n	Zero-mean gaussian noise
Σ_n	Noise covariance matrix
C_n	Normalized noise covariance matrix (voxel-independent)
σ_n	Noise variance (voxel-dependent)
h^*	best fit HDR when C_n is known
h	best fit HDR when the noise is white
$\hat{h}$	best fit HDR when C_n is estimated from residuals.
$\tilde{h}_{dt}$	best fit HDR with trend coefficients
e_{h^*}, $e_{\bar{h}}$, $e_{\hat{h}}$	estimation errors.
$\Sigma_{\hat{h}}$	Covariance matrix of $\hat{h}$
$C_{\hat{h}}$	Normalized covariance matrix of $\hat{h}$ (voxel-independent)
$\sigma_{\hat{h}}$	Variance of $\hat{h}$ (voxel-dependent)
R_e	Autocorrelation of residual errors
α,ρ	parameters used to fit R_e.
k_{max}	maximum number of coefficients of R_e to fit.
R	Restriction Matrix
q	Statistical test vector (equals Rh)
N_{tp}	Number of time points (ie, scans) per run.
N_h	Number of parameters (per stimulus) estimated in HDR.
N_c	Number of stimulus conditions.
N_{ch}	Total number of parameters estimated ($N_c * N_h$)
N_{dt}	Detrending order.
N_r	Number of runs.
N_q	Number of rows in the restriction matrix.
TR	Time between observations.
T_{HDR}	Time window over which to estimate the HDR ($N_h * TR$).
N_S	Number of sessions (or subjects)
N_R	Number of runs per session