Second-level (group) analysis methods

Theodore Huppert, PhD
(huppertt@upmc.edu)
University of Pittsburgh
Departments of Radiology and Bioengineering

- A look at general linear models
  - Whitening, coloring, and noise distributions
  - Canonical and deconvolution models
- Group-level models
  - Mixed Effects and ANOVA
  - Extension to image reconstruction
- Huppert Lab: NIRS-toolbox
  - Overview of toolbox structure
  - Example of GLM analysis
  - Example of group-level methods
  - Example of forward model and image reconstructions
### Second-level statistical models

#### AnalyzIR toolbox: Wilkinson-Rogers notation

<table>
<thead>
<tr>
<th>Formula</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta \sim -1 + \text{cond} + (1</td>
<td>\text{subject}) )</td>
</tr>
<tr>
<td>( \beta \sim -1 + \text{group:cond} + (1</td>
<td>\text{age}) )</td>
</tr>
<tr>
<td>( \beta \sim -1 + \text{group} + \text{cond} + \text{group*cond} + (1</td>
<td>\text{IQ}) )</td>
</tr>
</tbody>
</table>

\[
\beta = A \cdot \Gamma + B \cdot \Theta + \varepsilon
\]

where \( \beta \) is the vector of weights obtained from the first-level statistical model entries for each subject, task condition, and source-detector pair; \( A \) is the fixed effects model; and \( B \) is the random effects model matrices.

For example, this is using inclusion of the age as a cofactor:

\[
\begin{bmatrix}
\beta_{\text{Subj,CondLow}} \\
\beta_{\text{Subj,CondMed}} \\
\beta_{\text{Subj,CondHigh}} \\
\vdots \\
\beta_{\text{SubjN,CondLow}} \\
\beta_{\text{SubjN,CondMed}} \\
\beta_{\text{SubjN,CondHigh}}
\end{bmatrix} = \begin{bmatrix} 1 & 1 & \vdots & 1 & \vdots & 1 \end{bmatrix} \begin{bmatrix} \text{age}_A & \text{age}_A & \vdots & \text{age}_A \\
\text{age}_N & \text{age}_N & \vdots & \text{age}_N
\end{bmatrix} \begin{bmatrix} \Gamma_{\text{Low}} & \Gamma_{\text{Med}} & \vdots & \Gamma_{\text{High}} \\
\Gamma_{\text{Low.Age}} & \Gamma_{\text{Med.Age}} & \vdots & \Gamma_{\text{High.Age}} \\
\vdots & \vdots & \vdots & \vdots \\
\Gamma_{\text{Low.Age}} & \Gamma_{\text{Med.Age}} & \vdots & \Gamma_{\text{High.Age}}
\end{bmatrix} + \begin{bmatrix} 1 \\
1 \\
1 \\
\vdots \\
1 \\
1 \\
1
\end{bmatrix} \begin{bmatrix} \Theta_A \\
\Theta_M \\
\Theta_H \end{bmatrix} + \nu
\]

where the terms \( \Gamma_{\text{Low}}, \Gamma_{\text{Med}}, \) and \( \Gamma_{\text{High}} \) denote the main group level effects for the thee task conditions and terms \( \Gamma_{\text{Low.Age}}, \Gamma_{\text{Med.Age}}, \) and \( \Gamma_{\text{High.Age}} \) denote the interaction terms between the three conditions and age. The second matrix \( (B) \) and coefficients \( (\Theta) \) denote the random effects terms (here indicating subject as a random effect).
To solve the mixed effects

$$\Omega \cdot \beta = \Omega \cdot A \times \Gamma + \Omega \cdot B \cdot \Theta + \Omega \cdot \nu$$

where the whitening matrix is defined as

$$\Omega^T \Omega = \text{Cov}^{-1}_\beta$$

for reality (includes all source-to-detector channels simultaneously)

$$\Omega (A \otimes I_{\text{CHAN}}) \times \Gamma + \Omega (B \otimes I_{\text{CHAN}}) \Theta + \nu$$

where $I_{\text{CHAN}}$ is an identity matrix of size number of fNIRS source-to-detector pairs and $\otimes$ is the Kronecker operator.
Group-level image reconstruction

Measurement level: \[ Y_{\text{Subject}} = H \cdot \beta_{\text{Subject}} + u_{\text{Subject}} \]

Subject level: \[ \beta_{\text{Subject}} = \beta_{\text{Group}} + \Delta \beta_{\text{Subject}} \]

Group level: \[ \beta_{\text{Group}} = \beta_{0,\text{Group}} + \omega_{\text{Group}} \]

where \( \beta_{0,\text{Group}} \) is a prior on the expected value of the brain image for the group and the three noise terms \((u, \Delta \beta, \omega)\) are defined as follows (\( N_n(\mu, \Sigma) \) denotes \( N \)-variate normal distribution with mean vector \( \mu \) and covariance matrix \( \Sigma \) where \( N \) is the number of measurements):

\[ u_{\text{Subject}} \sim N_n(0, C_{\text{N}}^{-1}) \quad \Delta \beta_{\text{Subject}} \sim N_n(0, C_{\text{B}}^{-1}) \quad \omega_{\text{Group}} \sim N_n(0, C_{\text{G}}^{-1}) \]

The free-energy expression for the model

\[
\arg \max_{\beta, \omega, \Delta \beta} \left\{ -\frac{1}{2} \| Y - H \cdot \beta_{\text{Subject}} \|^2_{C_{\text{N}}} - \frac{1}{2} \| \Delta \beta_{\text{Subject}} \|^2_{C_{\text{B}}} - \frac{1}{2} \| \beta_{\text{Group}} - \beta_{0,\text{Group}} \|^2_{C_{\text{G}}} \right\}
\]

\[
\frac{1}{2} \log|C_{\text{N}}| - \frac{1}{2} \log|C_{\text{B}}| - \frac{1}{2} \log|C_{\text{G}}|
\]

GLM Coefficients Average Response

\[ \beta = L \cdot B + u \]

\[
\begin{array}{ccc}
\text{Subj1} & \beta_{1A} & A \\
\text{Subj1} & \beta_{1B} & B \\
\text{Subj2} & \beta_{2A} & A \\
\text{Subj2} & \beta_{2B} & B \\
\text{Subj3} & \beta_{3A} & A \\
\text{Subj3} & \beta_{3B} & B \\
\end{array}
\]

\[ \beta \sim -1 + \text{cond} \]

\[
\begin{bmatrix}
\beta_{1A} \\
\beta_{1B} \\
\beta_{2A} \\
\beta_{2B} \\
\beta_{3A} \\
\beta_{3B} \\
\beta_{1A} \\
\end{bmatrix} =
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
1 & 0 \\
0 & 1 \\
1 & 0 \\
0 & 1 \\
0 & 1 \\
\end{bmatrix}
\cdot
\begin{bmatrix}
\beta_{\text{cond},A} \\
\beta_{\text{cond},B}
\end{bmatrix}
\]
Similar assumptions

Noise ($\varepsilon$) is:
- normally distributed
- Homoscedasticity
- i.i.d.

Is this true for NIRS?
- Noise across subjects is not normal
- Heteroscedasticity  e.g. Bad contact in sensors
- Channels are not independent

So we should use the same pre-whitening and robust regression concepts as before

$$W \cdot \beta = S \cdot W \cdot L \cdot B + u \cdot w$$
and we already know the whitening from the GLM model

$$W = \frac{1}{\sqrt{\text{Cov}(\beta)}}$$

* $W = 1/\text{cholesky(Cov(\beta))}$

Does response vary with age?

$$\beta = -1 + \text{cond} + \text{cond:age}$$

With Subject as a random variable

$$\beta = -1 + \text{cond} + \text{cond:age} + (1|\text{Subject})$$

$$\beta = L \cdot B + Z \cdot A + u$$
Model for a single source-detector pair

\[ \beta = L \cdot B + Z \cdot \Gamma + u \]

Model for whole NIRS probe at once

\[ \beta = (L \otimes I_{\text{src-det}}) \cdot B + (Z \otimes I_{\text{src-det}}) \cdot \Gamma + u \]

Accounts for covariance between src-dets
Allows region-of-interest analysis

\[ T_{\text{RCV}} = \frac{c \cdot B}{\sqrt{c \cdot \text{COV}_B \cdot c^*}} \]

Model for image reconstruction

\[ \beta = (L \otimes A_{\text{DOT}}) \cdot B + (Z \otimes A_{\text{DOT}}) \cdot \Gamma + u \]

GLM Coefficients
(Src-Det)

Average Response
(Voxels)

- Estimation of image is ill-posed (non-unique solutions)
- Find solutions simultaneously consistent with all subjects
- NIRS probe may have blind spots (e.g. nearest neighbor probe)
- Fill in missing data from other subject's data
- Variations in probe placement
- Can use individual registrations
- Subject anatomy varies
- Can use individual head models for forward models
• **Differences from fMRI**
  – Few spatial dimensions/Lots of time points
    • c.f. 2D-REML methods used in SPM8
  – Non-uniform spatial noise
    • Bad coupling/per wavelength
  – (Too) Fast sample rate
    • Degrees-of-freedom? Independent measurements
  – Motion Artifacts
  – Multivariate (HbO2 v. Hb)
    • Definition of “activation”?

\[ Y = X \cdot \beta + \varepsilon \]

**Design Matrix**
- Canonical (e.g. fMRI)
- or FIR (e.g. HOMER-1/ERP*)
- Nuisance regressors (e.g. DCT)