Identifying Functional Co-Activation Patterns in Neuroimaging Studies Via Poisson Graphical Models

by: W Xue, J Kang, F DuBois Bowman, T Wager & J Guo

Biometrics 70(4), 812-822

January 30, 2015
fMRI experiments pipeline

A

fMRI Time Series -> Preprocessing -> Statistical Analysis

Peak Activation Locations – Foci

Statistical Parametric Maps (T maps)

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>-79</td>
<td>13</td>
</tr>
<tr>
<td>30</td>
<td>-83</td>
<td>23</td>
</tr>
<tr>
<td>18</td>
<td>-96</td>
<td>23</td>
</tr>
<tr>
<td>-16</td>
<td>-96</td>
<td>18</td>
</tr>
<tr>
<td>-22</td>
<td>-90</td>
<td>-17</td>
</tr>
<tr>
<td>-36</td>
<td>-93</td>
<td>14</td>
</tr>
</tbody>
</table>
# Foci from Different Contrasts

<table>
<thead>
<tr>
<th>Study</th>
<th>Contrast</th>
<th>Emotion</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>sad</td>
<td>-59</td>
<td>-14</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>happy</td>
<td>48</td>
<td>-78</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>anger</td>
<td>48</td>
<td>-74</td>
<td>-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>sad</td>
<td>-57</td>
<td>-13</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>disgust</td>
<td>35</td>
<td>-80</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# Regional Foci Count

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Region 1</th>
<th>Region 2</th>
<th>...</th>
<th>Region 19</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>...</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0</td>
<td>...</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>437</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

---

CBMA data & co-activation

Xue et al, 2014
The Bivariate Poisson Model

- \( i, j \) are 2 regions in brain, \( k \) denotes contrasts
- \( X_{i,k}, X_{j,k} \) the number of foci
- When \((X_{i,k}, X_{j,k}) \sim \text{BP}(\lambda_{ii}, \lambda_{jj}, \lambda_{ij})\) then:
  - \( X_{i,k} \sim \text{Pois}(\lambda_{ii} + \lambda_{ij}) \)
  - \( X_{j,k} \sim \text{Pois}(\lambda_{jj} + \lambda_{ij}) \)
  - Pmf is:

\[
P(X_{i,k} = x_{i,k}, X_{j,k} = x_{j,k}) = e^{-(\lambda_{ii} + \lambda_{jj} + \lambda_{ij})} \frac{\lambda_{ii}^{x_{i,k}} \lambda_{jj}^{x_{j,k}} \min(x_{i,k}, x_{j,k})}{x_{i,k}! x_{j,k}!} \sum_{s=0}^{\min(x_{i,k}, x_{j,k})} \binom{x_{i,k}}{s} \binom{x_{j,k}}{s} s! \left( \frac{\lambda_{ij}}{\lambda_{ii} \lambda_{jj}} \right)^s
\]
Network detection

- Parameter $\lambda = (\lambda_{ii}, \lambda_{jj}, \lambda_{ij})'$ fully describes network
- Covariance parameter $\lambda_{ij}$ controls strength of co-activation
- A penalized likelihood framework is adopted:
  $$\ell(\lambda; \mathbf{X}_i, \mathbf{X}_j) - \theta \lambda_{ij}$$
- $\theta$ imposes sparsity on the network
Latent variable representation

- The likelihood of the model is computationally demanding
- Let $Y_{ij,k}$ be the total co-activations
- An alternative representation is:

  $$X_{i,k} = Y_{ii,k} + Y_{ij,k}$$

  $$X_{j,k} = Y_{jj,k} + Y_{ij,k}$$

- The complete model likelihood is then:

  $$\ell_{\text{comp}}(\lambda; Y_{ij}, X_i, X_j) - \theta \lambda_{ij}$$

- Now, the EM of Karlis (2003) can be used for estimation
Bivariate EM

- **E-step:**

\[ Y_{ij,k}^{(t+1)} = E[Y_{ij,k} | X_{i,k}, X_{j,k}; \lambda^{(t)}] \]

\[ = \min(x_{i,k}, x_{j,k}) \sum_{y_{ij,k}=0}^{y_{ij,k}} \frac{y_{ij,k} P(Y_{ij,k}, X_{i,k}, X_{j,k}; \lambda^{(t)})}{\sum_{y_{ij,k}=0}^{y_{ij,k}} P(Y_{ij,k}, X_{i,k}, X_{j,k}; \lambda^{(t)})} \]

- **M-step:**

\[ \lambda_{ij}^{(t+1)} = \frac{\sum_{k=1}^{n} Y_{ij,k}^{(t+1)}}{\theta + n}, \]

\[ \lambda_{ll}^{(t+1)} = \frac{1}{n} \sum_{k=1}^{n} X_{l,k} - \frac{\theta + n}{n} \lambda_{ij}^{(t+1)} \quad \text{for } l = i, j. \]
The multivariate case

- Only 2-way interactions considered here
- Similar arguments as with 2D model:

\[ X_{i,k} = \sum_{j=1}^{p} Y_{ij,k} \quad i = 1 \ldots, p \]

- EM now minimizes:

\[
-l_{\text{comp}}(\lambda; \tilde{Y}_1, \ldots, \tilde{Y}_n, X_1, \ldots, X_n) + \theta \sum_{i=1}^{p} \sum_{j=i+1}^{p} \lambda_{ij} \\
= \sum_{k=1}^{n} \sum_{i=1}^{p} \sum_{j=i}^{p} [\lambda_{ij} - Y_{ij,k}\log(\lambda_{ij})] + \theta \sum_{i=1}^{p} \sum_{j=i+1}^{p} \lambda_{ij}.
\]
Choice of tuning parameter

- Optimal value of $\theta$ is not known in practice
- Idea: split data in $X_{\text{train}}, X_{\text{test}}$ and use predictive log-likelihood:

$$l_{\text{obs}}(\hat{\lambda}(\theta); X_{\text{test}}) = \sum_{k=1}^{n} l_{\text{obs}}(\hat{\lambda}(\theta); X_{\text{test},k}).$$

- $n$-fold cross validation for many $X_{\text{test}}$
Testing significance

- 2 tests for significance of findings (non-zero covariances)
- Model detects networks but tests build on $\lambda_{ij}^{\text{MLE}}$ (no penalization)
- Test I for pairs
  - $H_0 : \lambda_{ij} = 0$ vs $H_1 : \lambda_{ij} > 0$
  - Contrast labels permuted for $p$-values
  - FDR applied
- Test II for full functional networks $\Phi$
  - $H_0 : \lambda_{ij} = 0, \forall \{i,j\} \in \Phi$ vs $H_1 : \exists \{i,j\} \in \Phi : \lambda_{ij} > 0$
  - Same permutation procedure as before
Simulation studies 1/3

- Setup 1: 3 regions, 300 datasets, 100 bootstrap replicates for s.e.:

\[
\lambda = \begin{bmatrix}
1 & 3 & 1 \\
2 & 5 & \\
3 & \\
\end{bmatrix}
\]

- Setup 2: 8 regions, 500 datasets:

\[
\lambda_{12} = 3 \quad \lambda_{15} = 4 \quad \lambda_{16} = 2 \quad \lambda_{27} = 2 \\
\lambda_{36} = 3 \quad \lambda_{48} = 4 \quad \lambda_{57} = 5 \quad \lambda_{78} = 1
\]
### Penalized multivariate Poisson model

<table>
<thead>
<tr>
<th>Bias (%)</th>
<th>Coverage rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0020 (0.20%)</td>
<td>93.33% 95.00% 90.67%</td>
</tr>
<tr>
<td>0.0142 (0.71%)</td>
<td>96.00% 95.00%</td>
</tr>
<tr>
<td>0.0115 (1.15%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Covariance method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias (%)</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>0.1657 (16.57%)</td>
</tr>
<tr>
<td>0.0576 (2.88%)</td>
</tr>
<tr>
<td>0.0624 (2.08%)</td>
</tr>
<tr>
<td>0.1381 (13.81%)</td>
</tr>
<tr>
<td>0.1457 (2.91%)</td>
</tr>
<tr>
<td>0.2747 (9.16%)</td>
</tr>
</tbody>
</table>
Data

- 437 contrasts from 162 studies of emotion (Kober et al., 2008)
- On average, 6 foci per contrast
- GSK CIC atlas based on Harvard-Oxford atlas
- 19 ROIs in total
  - Dorsolateral prefrontal cortex reported w.p. 0.5 (highest)
  - Right globus pallidus reported w.p. 0.007 (lowest)
  - Others reported w.p. 0.140 (average)
Emotional processing network detected: 17 ROIs, 79 connections

Xue et al., 2014
Anterior cingulate cortex (ACC) most connections: 11

- Orbitofrontal cortex $\hat{\lambda}_{ij} = 0.023, p < 0.005$
- Striatum $\hat{\lambda}_{ij} = 0.018, p < 0.005$
- Thalamus $\hat{\lambda}_{ij} = 0.013, p < 0.005$

Overall network significant as well $p < 0.005$

- Clustering coef. $C = 0.710$, path length $L = 1.129$
- Several regions with high degrees:
  - Right insular $D = 14$
  - Thalamus $D = 14$
  - Left amygdala $D = 11$
Results 3/4

Motivation
Methods/Results
Conclusions

Xue et al, 2014

Neuroimaging stats group
Results 4/4

Neuroimaging stats group  Xue et al, 2014
Pros:
- New tool for CBMA data
- Likelihood based
- Nice interpretability

Cons:
- Brain tessellation may be subjective
- No 3-way (or more) interactions
- No voxel-wise rates

Xue et al, 2014
THANK YOU!!!