Hierarchical Linear Models for MEEG

Cyril Pernet,
University of Edinburgh, BRIC

Language and Categorization Lab.
http://www.sbirc.ed.ac.uk/lcl/
Overview

• Quick recap why we can do better than averaging
• Hierarchical Linear Modelling for MEEG
  ➢ What’s the hierarchy
  ➢ 1st level
  ➢ 2nd level
  ➢ MCC

• Examples
• Future Directions
Limitations of averaging
What’s wrong with current methods?

• Collect 132 channels and for each stimulus the neural time course from e.g. -100 to 1000ms

• Average all the trials of interest (e.g. 2 conditions) and look at where we have the strongest signal

• For each subject, either measure peak latency, peak amplitude or integrate portion of a component then do the stats across subjects
What’s wrong with current methods?

- Average all the trials of interest (e.g. 2 conditions) and look at where we have the strongest signal

- Why only looking at some electrodes? Why not analysing each electrode?

- Why averaging? Would be nice to account for the inter-trial variance (maybe there is something interesting in there). Can we miss something only looking the avg ERP?
What’s wrong with current methods?

• For each subject, either measure peak latency, peak amplitude or integrate portion of a component then do the stats across subjects

• Why only looking at peaks? Why not testing all of the signal?
What’s wrong with peaks anyway?

- Several lines of evidence suggest that peaks mark the end of a process and therefore it is likely that most of the interesting effects lie in the ‘component’ before a peak.

- **Neurophysiology**: whether ERPs are due to additional signal or to phase resetting effects a peak will mark a transition such as neurons returning to baseline, a new population of neurons increasing their firing rate, a population of neurons getting on / off synchrony.

- **Neurocognition**: reverse correlation techniques showed that e.g. the N170 component reflects the integration of visual facial features relevant to a task at hand (Schyns and Smith) and that the peak marks the end of this process.
Hierarchical Linear Modelling
What does LIMO EEG do?

• LIMO EEG allows analyzing
  ➢ all electrodes
  ➢ all time frames
  ➢ take into account the inter trial variability
  ➢ within + between variance = random effect
  ➢ Hierarchical modelling
What does LIMO EEG do?

- **LIMO EEG = Linear MOdelling of EEG data**
  
- Linear because we use parametric statistics which assume linear relationships between variables
  
- Modelling because we rely on the description of experimental effects (≠ model free approaches like ICA)
What does LIMO EEG do?

- Hierarchical Linear Modelling

   RTs of each subject
   
   Mean RTs of each subject (~level 1)
   -- here with a mean we lose within subject variance
   
   Mean RTs across subjects and std T-tests, ANOVA, etc.. (level 2)

   EEG of each subject
   
   Parameters for each subject (level 1)
   (all electrodes and all time points)
   
   Mean parameters across subjects and std T-tests, ANOVA, etc.. (level 2)
LIMO EEG: 1\textsuperscript{St} Level

\[ Y = X \ast B + e \]

All trials (1 time frame) (1 channel)  
Design matrix  
Coefficient to find  
Residuals

\[-2.2143; -2.5084; -1.8510; -4.7227]'

\[ Y = X \ast B + e \]

\[-2.2143; -2.5084; -1.8510; -4.7227]'

\[ Y = X \ast B + e \]
LIMO EEG: 1\textsuperscript{St} Level

\[ B(\text{est}) = \text{diag}(\text{pinv}(X^TX)X^TY) \]

All time frames \[ \equiv \]
Projection Matrix \[ \star \]
All trials / time frames
LIMO EEG: 1\textsuperscript{St} Level
LIMO EEG: 1\textsuperscript{st} Level

- **1st level**
- Assume data are normally distributed
- Obtain parameters using an OLS approach
- Analyze each time frame and electrode as independent (\textasciitilde assume residuals are iid)
- Since trials are ‘independent’ assumptions hold
LIMO EEG: 2\textsuperscript{St} Level

- 2\textsuperscript{nd level}

- Assume data are roughly normally distributed
- Estimate effect sizes and p values using bootstrap (residuals do not need to be iid)
- Analyze each time frame and electrode as independent then correct for multiple comparisons using maximum statistics / spatial-temporal clustering / temporal clustering
MCC in LIMO EEG

- Compute the same test under H0 (i.e., centre the data or resample in any order)
- Gives the probability to find clusters by chance
- Do it 1000 times → distributions of the cluster sums under H0 (sum of F values: reflect size and strength)
- Threshold observed clusters
MCC in LIMO EEG

Uncorrected $\rightarrow$ 132 electrodes and 201 times frames = 26532 tests

Alpha = 5% $\rightarrow$ 1327 false positives
LIMO EEG: Applications
LIMO EEG: Applications

• Examples from Rousselet et al.


LIMO EEG: Applications

• Subjects have to decide which of two faces is presented – phase coherence (noise) is manipulated across trials
LIMO EEG: Applications

- Investigate the effect of the stimuli properties on the neural responses

\[ EEG = \beta_1 + \beta_2 S + \beta_3 \phi + \beta_4 \gamma_2 + \beta_5 \gamma_1 + \beta_6 \phi \gamma_2 + \beta_7 \phi \gamma_1 + \varepsilon \]

- \( S \) stimulus identity (Face A vs. Face B)
- Phase (\( \phi \)), skewness (\( \gamma_1 \)), kurtosis (\( \gamma_2 \))
- \( \phi \gamma_2 \) phase-kurtosis interactions, \( \phi \gamma_1 \) phase-skewness interactions
- \( \varepsilon \) the error term

- Regression coefficients (\( \beta \)) are expressed in an arbitrary unit that reflects the strength of the fit (i.e. the influence of the factor on the EEG signal).
LIMO EEG: Applications
LIMO EEG: Applications

Partial correlations

Mean ERP amplitude (μV)

Original 100% phase coherence ERP
Explained variance
LIMO EEG: Applications

- Test for difference of sensitivity to noise in young vs. old subjects
LIMO EEG: Applications

- $R^2$ in regression reflects the % of variance explained by the model, here face A/B and global/local phase coherence.
LIMO EEG: Applications

• All designs can be analyzed

• Several conditions (ANOVA 1\textsuperscript{st} and 2\textsuperscript{nd} level)

• Several conditions but unwanted variance in the stimuli (ANCOVA 1\textsuperscript{st} level – ANOVA 2\textsuperscript{nd} level)

• One or many conditions with manipulation of interest in the stimuli (ANCOVA 1\textsuperscript{st} level Regression 2\textsuperscript{nd} level)
Future Directions
Robust Statistics

• Making 1st level inference robust
• Iterative Reweighted least square → weights relate to precision (inv. Var)
• Usually outliers are defined for a given Y but here we need to weight trials, i.e. find multivariate outliers
• Currently investigating Mid-Covariance Determinant
Building Models

- Model comparison over the whole space
- Investigate which model fits best the data
- The design matrix can reflect different theoretical models and/or sub-models
- Comparing models in time for a cluster
- Comparing models in space in different clusters?
- Comparing models in space and time?

General idea: Rousselet, Pernet 2011 Front. Psy
Collaborators:

Nicolas Chauveau (INSERM, France)
Carl Gaspar (CCNi, Psychology, Glasgow, UK)
Guillaume A Rousselet (CCNi, Psychology, Glasgow, UK)

LIMO EEG is freely available @
https://gforge.dcn.ed.ac.uk/gf/project/limo_eeg/