Signal Processing for Functional Brain Imaging: Summary

Dimitri Van De Ville
Medical Image Processing Lab, EPFL/UniGE
dimitri.vandeville@epfl.ch

Melissa Saenz
Laboratory of Neuroimaging Research, UNIL/EPFL
saenz.melissa@gmail.com

May 23, 2013
Types of data analysis

**Confirmatory**

*Does the model fit the data well?*

Problem → Data → Model → Analysis → Results

*Results depend on the model*

- GLM
- rt-FMRI

**Exploratory**

*Is there anything interesting in the data?*

Problem → Data → Analysis → Model → Results

*Can lead to unexpected results*

- unsupervised: ICA
- supervised: classification
Importance of understanding methods

- Dangers of circular analysis...

Can lead to bad publicity!...

Voodoo Correlation in Social Sciences

[Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition, Vul et al., 2009]
Examples of circular practices

- same data used for training and testing a classifier
- non-independent data used for training and testing
- all data used to define the ROI for a classifier analysis
- ROI-average activation regressed onto some factor that is related to the ROI-definition contrast
- descriptive or inferential analysis of ROI-average activation not independent of ROI definition
Example question 1

Parametric statistical testing generally uses more assumptions than non-parametric statistical testing (e.g., permutation testing).

☐ True     ☐ Not true

In general, the training phase of a classifier is computationally much less demanding than its application to new data (of the same size).

☐ True     ☐ Not true

Naïve Gaussian Bayes classifiers estimate the full covariance matrix of the data.

☐ True     ☐ Not true

Random linear mixing improves normality.

☐ True     ☐ Not true
Example question 2

Consider conventional statistical hypothesis testing for a general linear model (GLM), in the way it is performed by fMRI software such as SPM.

What does the by the statistical threshold parameter control? Explain briefly.

☐ Sensitivity ☐ Specificity

When the threshold parameter is exceeded, the null hypothesis (=contrast can be explained by noise) is rejected. This control specificity; i.e., the probability that the wrong decision was taken (=false positive).

Is this a good tool to be applied in neurosurgery? Why?

☐ Yes ☐ No

Controlling specificity (false positives) only relates to falsely declaring voxels to be active. However, the neurosurgeon wants to make sure that truely active voxels are detected, which relates to sensitivity (~false negatives)
Example question 3

We consider the equivalent of a two-sample t-test under the form of a general linear model (GLM) with 10 subjects; i.e., the design matrix consists of two regressors that model the respective groups.

\[ \mathbf{y} = \mathbf{X} \beta + \mathbf{e} \]; where \( \mathbf{X} \) is shown at the right.

1) What is the contrast vector that extracts the measure that is tested for the two-sample t-test?

\[ \mathbf{c}^T = [1, -1] \]

2) The two-sample t-test is signed; how would you perform an unsigned test?

The F-test can be used for that, either with the contrast \([1, -1]\) or \([-1, 1]\).
Example question 3

3) Derive the matrix that estimates the weights of the regressors

   a. Construct the normal equations by multiplying with the proper matrix (hint: need to obtain a square matrix at the right-hand side)

      \[ X^T y = X^T X \beta \]

   b. Now derive the solution for \( \beta \):

      \[ \beta = (X^T X)^{-1} X^T y = X^+ y \]

      What is the condition for this to work? How does this translate on the regressors?

      The square matrix \((X^T X)\) needs to be invertible. Since it is basically a correlation matrix of the regressors, that means that the regressors need to be linearly independent.

4) Draw or write the design matrix for a paired t-test for 10 subjects
Example question 4

Small-world networks exhibit two characteristic graph measures: small shortest-path length and large local clustering.

Explain a procedure that allows you to obtain a small-world network starting from a regular lattice.

With a given probability $p$, we can rewire the existing connections; i.e., we take a random edge, and replace one node of the edge with a randomly selected other node. By increasing the probability, we get from a regular to a small-world graph, and then into a completely random one.
Example question 6

What makes voxel-based classification in fMRI a difficult problem?

The large number of dimensions!

FMRI has many voxels to be considered (10,000 to 100,000) and the number of instances to learn from is limited (10 to 1,000 at most).

Therefore, it’s a high-dimensional learning problem (curse of dimensionality), which is difficult and prone to overfitting.
Retinotopic mapping stimuli are designed to induce a traveling wave of neural activity across the visual field maps of the visual cortex.

Why is this mapping method called "phase-encoded"?

The mapping stimuli periodically cycle through positions in visual space.

As a result, voxels in primary visual cortices have cyclical responses that peak at different times, depending on when the stimulus passes through the voxels' receptive field.

In other words, all signals will have the same magnitude responses, but different phases, and the phase encodes the location to which they respond.
Good luck!

- AAC137, Wednesday June 19, 8h15-11h15

- Don’t forget to check the “Summary of Techniques”