Advanced Signal Processing II  
*(aka Acronym Day)*

Latency Jitter and Woody Filters (acronym free)  
Inference Problems with Scalp Topography (acronym free)  
PCA  
ICA  
Removal of OCULAR artifacts with ICA (and lots of acronyms)  
BESA  
Simultaneous EEG with ICA and fMRI!

The Problem of Latency Jitter

➢ The averaging assumption of invariance in signal is not always warranted  
➢ Especially for the later endogenous components  
➢ To the extent that the signal varies from trial to trial, the average will produce potentially misleading results  
➢ Two common possibilities:  
  ➢ Smearing of components;  
  ➢ will underestimate amplitude of component (especially a problem if comparing groups, one group with more latency jitter)  
➢ Bimodal or multi-bumped components

The Solution

➢ The Woody Adaptive Filter (Woody, 1967)  
➢ Based on Cross-correlation  
  ➢ Assumptions less restrictive than averaging methods  
  ➢ Waveform (morphology) must be constant across trials  
  ➢ Latency need not be constant

Details

➢ Cross-correlational series  
  ➢ For two waveforms the correlation between each of them is computed  
  ➢ first with no lag in time  
    a1, a2, ..., an  
    b1, b2, ... bn  
  ➢ then with one lagged with respect to the other  
    a1, a2, ..., an-1  
    b2, b3, ... bn  
  ➢ A series of correlation values is obtained by progressively increasing the size of the lag

Announcements

➢ Papers:  
  ➢ You will received highly personal canned email acknowledgement that it was received  
  ➢ You will receive commented version via email once all papers are graded  
➢ Take home final due May 9 at noon (hardcopy in my mailbox).  
➢ Course Evals  
  ➢ 3x5s
The Basic Idea

Sine
Cosine
Cross-Correlation

See ... CrossCorr_Sin.Cos.m

More Details

- Can be used as a "template matching" procedure
- Compare running average with raw EEG epochs
- This is a method of single-trial signal detection:
  - First create a template: either predetermined (e.g., sine wave) or
    empirically determined (e.g., average)
  - Then calculate cross-correlational series between each raw EEG
    epoch and the template
  - If some maximum correlation achieved, conclude signal is present
  - If correlation not achieved conclude absent
  - This can also be used as a method of determining the latency of a
    component (by examining the trial-by-trial shifts), or of determining
    the variability in response for a given individual (again by examining
    the trial-by-trial shifts)

Woody’s Instantiation

- The Woody Adaptive Filter (Charles Woody, 1967) is a special case and
  application of cross correlational technique
- The term "adaptive" refers to the fact that the template is not established a priori,
  but generated and updated by an iterative procedure from the data themselves
- Procedure
  - Initial template is usually either a half cycle of a sine or triangle wave, or the
    unfiltered average of single trials
  - Cross-lagged correlations (or sometimes covariances) are then computed between
    each trial and the template typically over a limited range of samples (e.g., region of
    P300, not over "invariant" components)
  - Each trial is then shifted to align it with the template at the value which yields the
    maximum cross correlation (or covariance)
  - A new template is then generated by averaging together these time-shifted epochs
  - Procedure is repeated using this new average as the template
  - repeated until the maximal values of the cross correlation become stable
  - often, average cross-correlation value increment monitored; if Δ increases < .005 or
    .001, then stability achieved
- Some implementations, trials which do not reach a minimum criterion (e.g., .30-.50)
  are discarded from subsequent template construction and perhaps from
  subsequent analysis altogether

Woody Filtering Demo!

Validity

- Seems to do a fair job of improving signal extraction if a few iterations are used and if the
  original signal itself is singly peaked
- Wastell(1977) reports a decline in the validity of the procedure if numerous iterations are used
- Therefore, unlike averaging, Woody filtering can only improve signal-to-noise ratio over a definite
  limit
- Suggests also that Woody may not be the solution under conditions of very low signal-to-noise ratio
Dimensionality explosions!

32, 64, 128, 256!!!

Principal Components Analysis

- A method for reducing massive data sets
- See Handout for gory details

PCA (1): The Data matrix

Data Matrix above shows only one site – could have multiple sites by adding rows for each subject
This data matrix is for “temporal PCA” but one could transpose for “spatial PCA”

PCA (2): The Score matrix

These scores for each subject are optimally weighted composites of the original data, designed to capture as much variance as possible with as few scores as possible.
But for conceptual ease, imagine 5 scores: P1, N1, P2, N2, P3 amplitude

PCA (3): The Loading matrix
(to guess what components mean)

Spatial PCA on Sample Data
PCA (3b): The Loading Map
(for Spatial PCA)

Reminder: The ERP from which it derives

PCA Component 2

PCA (4): Reconstructing Data Matrix

- $D_{Nxn} \approx S_{Nxm} * L_{mxn}$
- This reconstructed Data matrix will differ slightly from the original Data matrix because not all $n$ components are used.
- To the extent that the $m$ components account for most of the variance in the original data set, the reconstructed data matrix will closely approximate the original data matrix.

PCA (4): Caveat Emptor

- PCA is a linear model; assumes the components sum together without interaction to produce the actual waveform
- Sources of variance are orthogonal; if two sources are highly correlated, may result in a composite PCA component reflecting both
- Component invariability in terms of latency jitter across subjects
  - PCA does not distinguish between variations in amplitude vs variations in latency
  - Especially a problem in comparing control vs pathological groups; pathological groups will typically be more variable
  - Allen & Collins unpublished simulation study:
  - Two groups: Control & Pathological
  - Identical waveforms for each group differed only in latency
  - The two groups differed significantly on three of four principal component scores
  - In other words, if one indiscriminately interprets these as amplitude or morphology differences, one would be WRONG!!

ICA … a “better” PCA?

- PCA finds orthogonal components
  - First PC accounts for most variance
  - Next PC accounts for most remaining variance
  - Components will have orthogonal scalp distributions
- ICA separates temporally independent components
  - Also known as blind source separation
  - May or may not correspond to brain “hotspots” but do represent functional brain networks
- See:
  - http://www.sccn.ucsd.edu/~scott/tutorial/icafaq.html
  - http://sccn.ucsd.edu/~arno/ (ICA for Dummies!)
ICA Decomposition

\[ x \rightarrow W \rightarrow Wx \rightarrow W^{-1} \]

- ERP Data
- Activations
- Maps
- ICA Components

ICA vs PCA

Principal component analysis vs Independent component analysis

EEG data are mixtures of source signals

Cocktail Party

ICA/EEG Assumptions

- Mixing is linear at electrodes
- Propagation delays are negligible
- Component time courses are independent
- Number of components ≤ number of channels.

From Tzyy-Ping Jung, presented at EEGLab Workshop, Nov 8, 2007
ICA: The Projection Map

ICA: Trial by Trial IC Projection to Pz

ICs as Artifacts!

"Clinical" vs Actuarial Approaches

“Clinical” vs Actuarial Approaches

- Human raters
  - Good source of possible algorithms
  - Lousy at reliably implementing them
    - Inter-rater
    - Intra-rater
- Actuarial methods
  - Always arrive at the same conclusion
  - Weight variables according to actual predictive power


ICs as Artifacts!

ADJUST:
An automatic EEG artifact detector based on the joint use of spatial and temporal features

Mignon, Jovicich, Bruzzone, & Buiatti, 2010

ICs as Artifacts!

MARA (Multiple Artifact Rejection Algorithm)

FASTER (Fully Automated Statistical Thresholding for EEG artifact Rejection)

SASICA (a tool for implementing these and more)…

Mognon, Jovicich, Bruzzone, & Buiatti, 2010

Chaumon et al., 2015
Rare Events

- Expected properties
- Four high amplitude events in otherwise low amplitude time courses
- High spatial/temporal locus measures

Chaumon et al., 2015

Eye blinks

- Features used
  - Spatial Average Difference (SAD)
  - Temporal Kurtosis (TK)
- Frontal distribution
- High power in delta frequency band

Mognon, Jovicich, Bruzzone, & Bulatti, 2010

Vertical Eye Movement

- Features used
  - Spatial Average Difference (SAD)
  - Maximum Epoch Variance (MEV)
- Frontal distribution similar to that of an eye blink

Mognon, Jovicich, Bruzzone, & Bulatti, 2010

Horizontal Eye Movement

- Features used
  - Spatial Eye Difference (SED)
  - Maximum Epoch Variance (MEV)
- Frontal distribution in anti-phase (one positive and one negative)

Mognon, Jovicich, Bruzzone, & Bulatti, 2010

Generic Discontinuities

- Features used
  - Generic Discontinuities Spatial Feature (GDSF)
  - Maximum Epoch Variance (MEV)
- Variable distribution
- Sudden amplitude fluctuations with no spatial preference
  - Could be present in as little as one or 2 trials, and limited to 1 channel
- In component data scroll weird activity in the trial plotted on the IC activity

Mognon, Jovicich, Bruzzone, & Bulatti, 2010
Neural Sources of EEG

Source Analysis

- BESA -- Brain Electrical Source Analysis
- This is a model-fitting procedure for estimating intracranial sources underlying ERPs
  - Estimate -- if model fits, then data are consistent with these sources; yet there is no unique solution
  - Not for ongoing EEG -- too many sources

BESA

- Imagine a data matrix of ERPs: $V_{Cxn}$ (# Channels by # timepoints)
- Note that this is really the result of the subtraction of the activity at the reference from the activity at the these sites; i.e.,
  $$V_{Cxn} = U_{Cxn} - R_{Cxn}$$
- Note: the reference matrix has identical rows! Thus BESA Presumes that all channels referenced to the same reference!
Reconstruct a data matrix that includes not only the original channels, but the implicit channel (reference) as well:

\[ U_{Exn} (\# \text{electrodes} = \# \text{channels}+1) \]

which represents the activity at each electrode with respect to an average reference (i.e., the average of all channels).

Now this matrix \( U_{Exn} \) can be decomposed into:

- a set of sources: \( S_{xn} \) (# Sources by # timepoints)
- a set of attenuation coefficients \( C_{ExS} \)
- so that \( U_{Exn} = C_{ExS} S_{xn} \)

The attenuation matrix is determined by:

- the geometry between the source and the electrodes
- the nature of the conductance of the three-layer head model (Brain, Skull, Scalp);
  - the skull is less conductive than the layers on either side
  - this results in a spatial smearing of potentials as they cross the skull
  - the skull produces the equivalent of a brain that is 60% of the radius of the outer scalp (rather than the "true" figure of ~84%)

Note that the decomposition of \( U \) into \( C \) and \( S \) results in:

- an electroanatomical time-independent matrix (\( C \)) that reflects that anatomical substrates do not move around in the head
- a time-variant dipole source potential matrix that represents the change in activity of each source over time
BESA Vs PCA Vs ICA
(a battle of acronyms)

- This decomposition is akin to PCA/ICA
  - PCA and ICA have sources and propagation coefficients
  - PCA solutions are constrained by orthogonality of components, and by those that account for greatest common variance
  - ICA constrained to find temporally independent components
  - BESA solutions are constrained by the geometry of the head, the volume conduction of the dipoles, and the anatomical constraints dictated by the user (e.g., inside the head, symmetrical, not in the ventricles, must not be in the brainstem after a certain point in time, etc...)

BESA Vs PCA Vs ICA continued

- Like PCA/ICA, the reconstruction of the original data set will be imperfect
  - With all methods, better chance of reconstructing the original matrix if data are reliable
  - If you capture the important sources, the reconstruction should be very good (i.e., small residual variance)
  - It is useful to attempt to upset a solution by inserting another source and seeing if:
    - the original solution is stable
    - the new source accounts for any substantial variance
- Can do dipole localization (BESA) on an IC!

Dipole Fitting

PCA

ICA

You can try it!

Implementations

- BESA can be used:
  - in a strict hypothesis-testing manner by designating sources a priori and testing the fit
  - in an exploratory/optimizing manner by allowing the program to iteratively minimize the residual variance (between observed and reconstructed waveforms) by:
    - moving dipoles
    - changing the orientation of dipoles
    - altering the time-by-activity function of the dipoles
BESA – Did it work?

- In the end, the adequacy of your solution will be judged by
- stability of your solution:
  - against insertion of additional dipoles
  - across multiple subjects
- anatomical feasibility
- follow-up tests with patients with lesions
- your reviewers!

Recording EEG in fMRI environments: Oodles of Issues

- EEG can be bad for fMRI
- Wires and electrodes can be ferromagnetic = TROUBLE
- Wires and electrodes can be paramagnetic = less trouble
- MRI and fMRI can be bad for EEG
- Gradient switching creates huge artifact for EEG
- Movement in Magnetic fields creates current in any conductive medium (e.g. wires!)
- High frequency current can make wires HOT and RF is 127.68 MHz at 3T – that’s fast, and can create mega-hurts!
- Thus in-line 10K resistor

Special Caps

- Need conductive material
- That will not heat up
- That will not pose hazard in strong magnetic field
- That includes inline resistor to prevent any induced current from reaching the subject
- That includes Styrofoam head at no charge

Whence EEG Artifacts in fMRI?

- Faraday’s law of induction...
  - induced electromotive force is proportional to the time derivative of the magnetic flux
  - Flux = summation of the magnetic field perpendicular to the circuit plane over the area circuit
  - \( \varepsilon = \frac{d\Phi}{dt} \)
- Can reflect:
  - changes in the field
  - Changes in the circuit relative to the field
Whence EEG Artifacts in fMRI?

- **RF pulses**
  - For 3T = 127.6 MHz
  - Brain oscillations = 0.5-50 Hz
  - Amplifier frequency range = DC-3.0 KHz
  - Artifacts thus attenuated, but still range overwhelm the EEG signal

- **Gradient Switching**
  - Artifact approximates differential waveform of the gradient pulse
  - Polarity and amplitude varies across channels
  - Frequency = 500-900 Hz
  - EEG dominated by harmonics of slice repetition frequency (≈10-25 Hz)
  - convolved with harmonics of volume repetition frequency (≈0.2-2 Hz)
  - Artifacts in range from 1000-10,000 μV!

---

**Average Artifact (across 1 TR)**

Ritter et al., 2009

**RF = radiofrequency wave; Gs = slice selection gradient; Gp = phase encoding gradient; Gr = readout gradient; a = Fat suppression pulses (1-3-1-3-1 pulses); b = slice selection RF; c, d, h = spoilers; e = slice selection gradient; f = dephasing and rephasing gradient; g = readout gradient.‘ = EEG artifact corresponding to letter.**
Whence EEG Artifacts in fMRI?

- Faraday’s law of induction...
  - induced electromotive force is proportional to the time derivative of the magnetic flux
  - Flux = summation of the magnetic field perpendicular to the circuit plane over the area circuit
  - $\varepsilon = \frac{d\Phi}{dt}$

Can reflect:
- changes in the field (gradient switching, RF)
- Changes in the circuit geometry or position relative to the field due to body motion

Debener et al., 2008

EEG in Magnet (no scanning)

Simulated EKG Artifact

Axial rotation - low frequency spatially-distributed effect, with polarity reversal

Debener et al., 2009
REMOVING THOSE PESKY ARTIFACTS!

FASTR: FMRI Artifact Slice Template Removal

- Part of FMRIB Plug-in for EEGLAB
- Upsample to at least 20K Hz
- Align slices for slight jitter in timing
- Moving Window approach with subtraction
- PCA on artifact residuals form Optimum Basis Set (OBS) to reduce residual artifacts by 90%
- Downsample to original rate
- Sample Results

ECG-related removal via moving average subtraction (Allen et al. 1998)

Gradiant/RF removal via moving average subtraction
There may be residual crud (RC)

Simultaneous EEG and RS-fMRI (following ICA!)

Debener, Ullsperger et al. J Neurosci 2006

Multi-modal Imaging

Tether EEG asymmetry to other measures neural systems known to be involved in MDD

Mayberg et al., 2005

Multi-modal Imaging

Create RS-fMRI network with ACC seeds

Allen, Hewig, Mittner, Hecht, & Schnyer, In preparation

EEG Alpha Asymmetry is Negatively Correlated with IFG Connectivity in Two ACC-seeded Resting State Networks

Spatially-enhanced EEG asymmetry (using CSD transform) at sites F8-F7 is related to resting state connectivity between left inferior frontal gyrus and two ACC-seeded networks.

Allen, Hewig, Mittner, Hecht, & Schnyer, In preparation
EEG-fMRI Synopsis

Less relative left frontal activity (indexed by EEG) is related to increased connectivity of left IFG to two ACC-seeded RS networks

Consistent with:
- Hyper-connectivity in RSfMRI emotion networks in MDD (e.g., Grecius et al., 2007; Sheline et al., 2010)
- Frontal EEG asymmetry findings of less relative left frontal activity in risk for MDD.

Alpha power may regulate network connectivity

Note: Between vs Within Subjects

BETWEEN-SUBJECTS’ DATA DOES NOT NECESSARILY SUPPORT A WITHIN-SUBJECTS’ INTERPRETATION

Within Subjects’ Moderation of RSfMRI Connectivity

Calculate F8-F7 alpha asymmetry for each TR
- EEG leads TR by 4.096 seconds
- Median split into high (left) and low (right)
- Entered as moderator in PPI approach (cf. Friston et al., 1997)
- Tests whether strength of connectivity to seed region varies as a function of the moderator

Within Subjects’ Moderation of RSfMRI Connectivity

Dorsal ACC Seed

Greater Connectivity with Less Left Frontal Alpha or Greater Left Frontal Alpha

Cognitive Control over Emotion

IFG has a key role in mediating the success of cognitive control over emotional stimuli
Cognitive Control over Emotion

Left IFG:
Language and self-referential processing

Right IFG:
Attentional control
behavioral inhibition
suppression of unwanted thoughts
attention shifting
efforts to reappraise emotional stimuli

Working Hypothesis:
Hyperconnected left IFG and emotion networks: rumination
Hypoconnected right IFG: difficulty disengaging from emotion

Psychophysiology -- Synopsis

Psychophysiology is inherently interdisciplinary, and systemic
Principles learned here can apply to a wide range of physiological signals
- Recording
- Processing
- Interpretation

Mundane Details

Exams due Monday May 9 by noon in my mailbox, room 312 Psychology.

Papers will be emailed to you

Final grades will be available for lookup on the web; email will alert you