SPATIO-TEMPORAL MODELS WITH APPLICATIONS TO HEALTH DATA

Hernando Ombao

University of California at Irvine
Department of Statistics

December 5, 2012
Outline of Talk

1. Research Motivation
2. Vector Autoregressive Models
3. Mixed Effects VAR
4. Spectral Dependence
5. Biomarker Selection
Outline of Talk

1. Research Motivation
2. Vector Autoregressive Models
3. Mixed Effects VAR
4. Spectral Dependence
5. Biomarker Selection
Outline of Talk

1. Research Motivation
2. Vector Autoregressive Models
3. Mixed Effects VAR
4. Spectral Dependence
5. Biomarker Selection
OUTLINE OF TALK

1. Research Motivation
2. Vector Autoregressive Models
3. Mixed Effects VAR
4. Spectral Dependence
5. Biomarker Selection
1. Research Motivation
2. Vector Autoregressive Models
3. Mixed Effects VAR
4. Spectral Dependence
5. Biomarker Selection
**Research Motivation**

- Visual-motor electroencephalogram (HAND EEG)
- Motor-decision (fMRI)
- LA county environmental data (mortality, pollution, temperature)
Electrophysiologic data: multi-channel EEG, local field potentials
Hemodynamic data: fMRI time series at several ROIs
Multi-channel (multivariate)
Two movement conditions: leftward vs. rightward
**Neuroscience Data and Statistical Goals**

- **External Stimulus**
  - Visual, Auditory, Somatosensory, Stress
- **Personality traits, Genes, Socio-Environmental Factors**
- **Unobserved**: brain network/cell assemblies
- **Brain Signals** (indirect measures of neuronal activity)
  - Functional: fMRI, EEG, MEG, PET
  - Anatomical: DTI
- **Acute Outcomes**
  - Emotion, Skin conductance, Motor response
**Neuroscience Data and Statistical Goals**

- Stimulus
- Neuronal Response
- Brain Signals
- Behavior
NEUROSCIENCE DATA AND STATISTICAL GOALS

Stimulus  →  Neuronal Response  →  Brain Signals  →  Behavior

Moderators Modifiers

Genes Trait Socio-Environment
NEUROSCIENCE DATA AND STATISTICAL GOALS

Changes in the mean

Changes in variance

Changes in Cross-Dependence
Our Research Goals

- Characterize dependence in a brain network
  - Temporal: $Y_1(t) \sim [Y_1(t - 1), Y_2(t - 1), \ldots]'$
  - Spectral: interactions between oscillatory activities at $Y_1, Y_2$
- Develop estimation and inference methods for connectivity
- Investigate potential for connectivity as a biomarker
  - Predicting behavior
    - Motor intent (left vs. right movement)
      [Brain-Computer-Interface]
    - State of learning
    - Level of mental fatigue
  - Differentiating patient groups (bipolar vs. healthy children)
    - Connectivity between left DLPFC $\leftrightarrow$ right STG is greater for bipolar than healthy
Our Research Goals

- New dependence measures must be easily interpretable
- Models must incorporate information across trials, across subjects
- Models must account differences in brain network between conditions
- Take advantage of multi-modal data (EEG, fMRI, DTI)
  - Model should be informed by physiology and physics
- Dimension reduction: extract information from massive data that is most relevant for estimating dependence
- Develop formal statistical inference procedures
Our Research Goals

Selected Contributions to the Time Series Literature

- **Automatic methods:**
  - SLEX Transform (**Smooth Localized Complex EXponentials**)
    - Ombao et al. (2001, JASA)
    - Ombao et al. (2001, Biometrika)
    - Ombao et al. (2002, Ann Inst Stat Math)
    - Huang, Ombao and Stoffer (2004, JASA)
    - Ombao et al. (2005, JASA)
    - Böhm, Ombao et al. (2010, JSPI)

- **Massive data; Complex-dependence; Mixed Effects**
  - Freyermuth, Ombao, von Sachs (2009, JASA)
  - Gorrostieta, Ombao et al. (2012, NeuroImage)
  - Kang, Ombao et al. (2012, JASA)
VAR MODELS AND APPLICATION TO THE LA COUNTY MORTALITY DATA

Cardiovascular Mortality

Temperature

Particulates
VAR Models and Application to the LA County Mortality Data

- Shumway, Azari and Pawitan (1988); Shumway and Stoffer (2010)
- LA County
- Weekly data on mortality, temperature and pollution levels
- Model mortality \( \sim \) (temperature + pollution)
- Granger causality: does past knowledge of temperature and pollution help improve prediction for mortality?
- Causation vs Association
- Practical issues
  - Hospitalization (rather than mortality)
  - Effect of pollution might be long term (rather than short term)
VAR Models and Application to the LA County Mortality Data

The VAR Model

- \( Y(t) = [Y_1(t), Y_2(t), Y_3(t)]' \)
- \( Y(t) = [\text{Mort}(t), \text{Temp}(t), \text{Part}(t)]' \)
- VAR(1) Model

\[
\begin{align*}
Y_1(t) &= \phi_{11} Y_1(t - 1) + \phi_{12} Y_2(t - 1) + \phi_{13} Y_3(t - 1) + \epsilon_1(t) \\
Y_2(t) &= \phi_{21} Y_1(t - 1) + \phi_{22} Y_2(t - 1) + \phi_{23} Y_3(t - 1) + \epsilon_2(t) \\
Y_3(t) &= \phi_{31} Y_1(t - 1) + \phi_{32} Y_2(t - 1) + \phi_{33} Y_3(t - 1) + \epsilon_3(t)
\end{align*}
\]

- In matrix notation VAR(1)

\[
Y(t) = \Phi Y(t - 1) + Z(t)
\]
VAR Models and Application to the LA County Mortality Data

Focus only on the dynamics of cardiac mortality

- Trend $\mu(t)$ - linear trend + seasonality
- The lag-1 model

\[
\text{Mort}(t) = \mu(t) + \phi_{11}\text{Mort}(t-1) + \\
\phi_{12}\text{Temp}(t-1) + \phi_{13}\text{Part}(t-1) + Z_1(t)
\]

- Lagged dependence parameters
  - $\phi_{11}$: $\text{Mort}(t-1) \rightarrow \text{Mort}(t)$
  - $\phi_{12}$: $\text{Temp}(t-1) \rightarrow \text{Mort}(t)$
  - $\phi_{13}$: $\text{Part}(t-1) \rightarrow \text{Mort}(t)$
VAR Models and Application to the LA County Mortality Data

Results.

- Bayesian information criterion chose $L = 2$ as the optimal lag
- Predicted cardiac mortality at time $t$

$$\hat{M}(t) = 56 - 0.01t + 0.30M(t - 1) - 0.20T(t - 1) + 0.04P(t - 1) + 0.28M(t - 2) - 0.08T(t - 2) + 0.07P(t - 2)$$
MIXED EFFECTS VAR AND APPLICATIONS TO BRAIN SIGNALS
Motor-Decision Experiment

- \( N = 15 \) right-handed college students
- Experiment: subjects see visual targets and must move joystick
- Two Conditions
  - Free choice - subject freely chooses any target
  - Instructed - subject must choose the specified target
- Regions of interest (7 areas that show highest differential activation)
  - PFC, SMA, etc.
Mixed Effects VAR and Applications to Brain Signals

- Limitations of the classical VAR model
  - Dependence and connectivity identical for all participants/subjects
  - Dependence and connectivity identical across all experimental conditions

- Our novel contribution: mixed effects VAR model
  - Gorrostieta, Ombao, et al. (2012, NeuroImage)
  - Subjects allowed to have a unique brain network
  - Model captures the effect of an experimental condition on connectivity
Mixed Effects VAR and Applications to Brain Signals

- Total of $R$ regions of interest (ROIs)
- $Y_r^n(t)$ the fMRI time series at the ROI $r$ for subject $n$
- Entire network: $Y^n(t) = [Y_1^n(t), \ldots, Y_R^n(t)]'$.
- General additive model

$$Y^n(t) = F^n(t) + E^n(t)$$

- The components
  - $F^n(t)$ - deterministic component
  - $E^n(t)$ - stochastic component
The deterministic component $F^n(t)$

- Decomposition

$$F^n(t) = D^n(t) + M^n(t) + \beta^n_1 \otimes X_1(t) + \ldots + \beta^n_C \otimes X_C(t),$$

- The mean component includes systematic changes in the BOLD signal that is due to
  - Scanner drift
  - Physiological signals of non-interest (e.g., cardiac and respiratory)
  - Experimental conditions
The stochastic component $E^n(t)$

- $E^n(t)$ captures between-ROI connectivity

$$\text{Cov}[Y^n(t + h), Y^n(t)] = \text{Cov}[F^n(t + h) + E^n(t + h), F^n(t) + E^n(t)]$$
$$= \text{Cov}[E^n(t + h), E^n(t)]$$

- $E^n(t)$ cannot be observed directly; we use the residuals:

$$E^n(t) = Y^n(t) - F^n(t)$$
$$R^n(t) = Y^n(t) - \hat{F}^n(t)$$
The ME-VAR(1) Model

$$E^{(n)}(t) = \left[ \Phi_{1,k} W_1(t) + \Phi_{2,k} W_2(t) + b_1^{(n)} \right] E^{(n)}(t - 1) + e^{(n)}(t)$$

- $W_1(t)$ is the indicator function
  - When condition 1 is active then $W_1(t) = 1$ and $W_2(t) = 0$
  - When condition 2 is active then $W_1(t) = 0$ and $W_2(t) = 1$
- $b_1^{(n)}$ models between-subject variation in connectivity
  - $b_1^{(n)} \sim (0, \sigma_1^2)$
Mixed Effects VAR and Applications to Brain Signals

- Subject-specific connectivity matrix (condition on $b_1^{(n)}$)
  - When $W_1(t) = 1$, the connectivity matrix for subject $n$ is
    \[ \Phi_{1,1} + b_1^{(n)} \]
  - When $W_2(t) = 1$, the connectivity matrix for subject $n$ is
    \[ \Phi_{1,2} + b_1^{(n)} \]

- The model can be utilized to test for
  - Lagged dependence between each pair of ROIs
    \[ H_0 : \Phi_{1,1} = 0, \Phi_{1,2} = 0 \]
  - Granger causality in each experimental condition
  - Testing for differences between conditions
    \[ H_0 : \Delta_1 = \Phi_{1,1} - \Phi_{1,2} = 0 \]
Mixed Effects VAR and Applications to Brain Signals
MIXED EFFECTS VAR AND APPLICATIONS TO BRAIN SIGNALS
Spectral Measures of Dependence: Partial Coherence

- Time series: $X, Y, Z$
- Cross-correlation $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}X \text{Var}Y}}$
- Partial cross-correlation between $X$ and $Y$ given $Z$
  - Remove $Z$ from $X$: $\epsilon_X = X - \beta_X Z$
  - Remove $Z$ from $Y$: $\epsilon_Y = Y - \beta_Y Z$
  - $\rho(X, Y|Z) = \frac{\text{Cov}(\epsilon_X, \epsilon_Y)}{\sqrt{\text{Var} \epsilon_X \text{Var} \epsilon_Y}}$
Spectral Measures of Dependence: Partial Coherence

<table>
<thead>
<tr>
<th></th>
<th>Model A</th>
<th>Model B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross-Corr</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Partial CC</td>
<td>NO</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Spectral Measures of Dependence: Partial Coherence

- Let $\mathbf{U}(t) = [X(t), Y(t), \text{others}(t)]'$.  
- Estimate partial coherence when $\text{dim} \mathbf{U}(t)$ is large.  
- Characterization of partial coherence (Dahlhaus, 1996)
  - Spectral matrix $\mathbf{f}(\omega)$
  - $\Lambda(\omega) = H(\omega) \mathbf{f}^{-1}(\omega) H(\omega)$
  - Partial coherence between $X$ and $Y$ is $|\Lambda_{12}(\omega)|^2$
- A Problem: bias in sample eigenvalues leads to poor condition number
  - Sample max eigenvalue over-estimates
  - Sample min eigenvalue under-estimates
Spectral Measures of Dependence: Partial Coherence

- Standard methods (Welch’s periodogram; multi-taper) tend to produce highly erratic results
- Our novel contribution: Shrinkage Method
  - Target: highly structured spectral matrix $\mathbf{V}(\omega)$
    - Vector auto-regressive; Vector ARCH
    - Diagonal matrix
  - Initial estimator $\tilde{I}(\omega)$ (non-parametric)
  - Generalized shrinkage estimator $\hat{f}(\omega) = (1 - W(\omega))\tilde{I}(\omega) + W(\omega)\mathbf{V}(\omega)$
    - $W(\omega) \propto \mathbb{E}[\tilde{I}(\omega) - f(\omega)]^2$
**Biomarker Selection**

- **Outcome of interest**
  - Behavioral measures (cognitive assessments, etc)
  - Response to treatment

- **Potential predictors**
  - Neuroimaging-derived measures
  - Clinical
  - Demographic
  - Genetic

- **BIG Problem**: The number of potential predictors exceed the number of subjects

- "Large $P$ - small $N$ problem"
**Biomarker Selection**

- **Regression Model**
  \[ Y_n = \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} + \ldots + \beta_P x_{Pn} + \epsilon_n \]

- **Least Squares Criterion**
  \[ C(\beta) = \sum_{n=1}^{N} [Y_n - (\beta_0 + \beta_1 x_{1n} + \ldots + \beta_P x_{Pn})]^2 \]

- **Penalty for complexity**
  \[ L_1(\beta) = \sum_{p=1}^{P} |\beta_p| \]
  \[ L_2(\beta) = \sum_{p=1}^{P} |\beta_p|^2 \]

- **Complexity-penalized least squares criterion**
  \[ PC(\beta) = C(\beta) + \lambda_1 L_1(\beta) + \lambda_2 L_2(\beta) \]

- **Result:** many \( \beta \) estimates will be forced to 0 (considered irrelevant!)
Biomarker Selection

- Complexity-penalized methods
  - LASSO, elastic net
  - Limitation: does not assess uncertainty in biomarker selection

- Our novel contribution: **Bootstrap-enhanced elastic net method**
  - Bunea, She, Ombao et al. (2011, NeuroImage)
  - Obtain $B$ bootstrap datasets
  - For each $b = 1 : B$ bootstrap data, record the predictors that were selected
### Biomarker Selection

#### Variable Selection

<table>
<thead>
<tr>
<th>Variable</th>
<th>Animal_T</th>
<th>COWAT_T</th>
<th>Trail_B_T</th>
<th>Trail_A_T</th>
<th>GPe_nondom_T</th>
<th>GPe_dom_T</th>
<th>WAIS_SymSrch_T</th>
<th>WAIS_DigSym_T</th>
<th>WAIS_LNS_T</th>
<th>BVMT_delay_T</th>
<th>BVMT_sum_T</th>
<th>HVLT_delay_T</th>
<th>HVLT_sum_T</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kmrk_cocopi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hcv_current</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kmrk_alc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fa_ic4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Vector (multivariate) Autoregressive Models
Mixed Effects VAR
Spectral Dependence: Shrinkage procedure
Biomarker selection: bootstrap enhanced elastic net
## US-Based Statisticians

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-chi Aban</td>
<td>Univ Alabama - Birmingham</td>
<td>Survival, Trials Cardiology</td>
</tr>
<tr>
<td>Dexter Cahoy</td>
<td>Louisiana Tech</td>
<td>Applied Stats</td>
</tr>
<tr>
<td>Mark Fiecas</td>
<td>Univ California San Diego</td>
<td>Applied Stats Imaging</td>
</tr>
<tr>
<td>Hernando Ombao</td>
<td>Univ California - Irvine</td>
<td>Time Series Brain Signals</td>
</tr>
<tr>
<td>Edsel Pena</td>
<td>Univ South Carolina</td>
<td>Survival, Nonpar</td>
</tr>
<tr>
<td>John Yap</td>
<td>FDA</td>
<td>Genetics</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENT

- INSTAT, UP Los Baños (Prof Reaño)
- Clinical Epidemiology, UP Manila