Penalized Subgrouping of Heterogeneous Time Series

Christopher Crawford, Jonathan Park, Sy-Miin Chow, & Zachary Fisher
The Pennsylvania State University
Department of Human Development and Family Studies
cmcrawford@psu.edu

June 28, 2023

Goals

- Contextualizing Heterogeneity
- Multi-VAR
  - Goals of Multi-VAR
  - Limitations
- Multi-VAR with Subgrouping
  - Why Subgroup?
  - Simulation Results
- Next Steps

Outline

1 Heterogeneity

2 Multi-VAR

3 Subgrouping Multi-VAR

4 Simulation Results

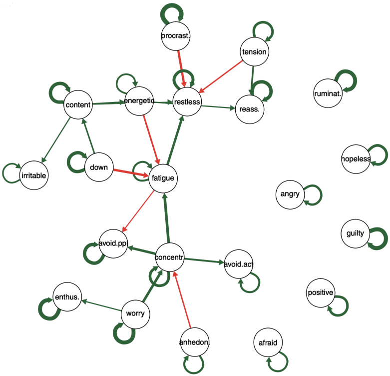
5 Future Directions

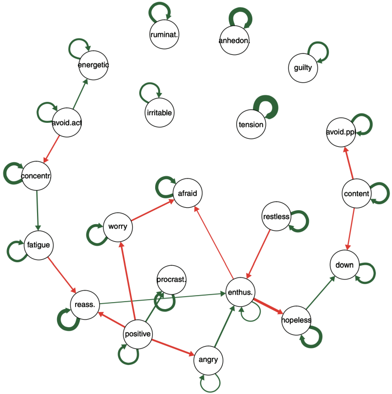
Types of Heterogeneity

- **Quantitative**: Differences in magnitude
- **Qualitative**: Differences in structure

- Major Depressive Disorder, for example, is characterized by both types of heterogeneity
  - Two individuals with a DSM-5 diagnosis of MDD could share no single symptom
  - Symptomatology can differ in both *presence* and *degree*

Depression Networks





Accounting for Heterogeneity

- Two ends of the heterogeneity spectrum: Nomothetic vs. Idiographic
- Multilevel models (and other flavors)
    - Account for heterogeneity in the *magnitude* of parameters
    - May be overly restrictive with respect to the functional form of the model
- Person-specific models
    - Allows for maximal flexibility
    - Disregards potential shared information
- Can we find a balance between these two?

Outline

1. Heterogeneity
2. Multi-VAR
3. Subgrouping Multi-VAR
4. Simulation Results
5. Future Directions

Multi-VAR

- Goal: Retain desirable features of canonical VAR while addressing concerns
- For multiple-subject ILD, we want to estimate $\mathbf{\Phi}^1, \ldots, \mathbf{\Phi}^K$ transition matrices for $1, \ldots, K$ individuals
- Consider the following decomposition of $\mathbf{\Phi}$:

$$\mathbf{\Phi}^k = \mathbf{\Gamma}^0 + \mathbf{\Gamma}^k, \; k = 1, \ldots, K$$

- $\mathbf{\Gamma}^0$ is a $d \times d$ matrix of common effects across K individuals
- $\mathbf{\Gamma}^k$ is a $d \times d$ matrix of effects unique to individual k
- Quantitative and qualitative differences across individuals

Standard Multi-VAR

- Fisher et al. (2022) proposed one approach for estimating $\boldsymbol{\Phi}^k$ using the Lasso penalty:

$$\underset{\boldsymbol{\Gamma}=(\boldsymbol{\Gamma}^0,\boldsymbol{\Gamma}^1,\ldots,\boldsymbol{\Gamma}^K)}{\operatorname{argmin}} \frac{1}{N}\sum_{k=1}^{K}\|\mathbf{Y}^k-(\boldsymbol{\Gamma}^0+\boldsymbol{\Gamma}^k)\mathbf{Z}^k\|_2^2+\lambda_1\|\boldsymbol{\Gamma}^0\|_1+\sum_{k=1}^{K}\lambda_2\|\boldsymbol{\Gamma}^k\|_1$$

- Sparsity in individual transition matrices $\boldsymbol{\Phi}^k$ induced and determined by penalty parameters λ_1 and $\lambda_{2,k}$
- Heterogeneity of solution determined by the competition of the two penalty parameters
- However, Lasso penalty suffers from a number of known issues (e.g., (Zhao & Yu, 2006)

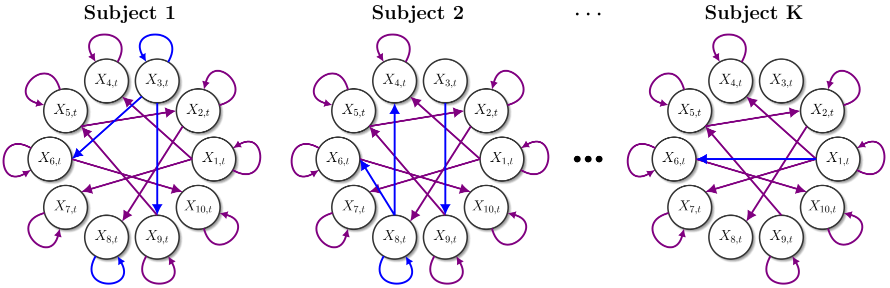
Adaptive Multi-VAR

- Fisher et al. (2022) proposed an objective function for Multi-VAR with adaptive Lasso (Zou, 2006):

$$\frac{1}{N}\sum_{k=1}^{K}\|\mathbf{Y}^k - (\mathbf{\Gamma}^0 + \mathbf{\Gamma}^k)\mathbf{Z}^k\|_2^2 + \lambda_1\boldsymbol{\omega}\|\mathbf{\Gamma}^0\|_1 + \sum_{k=1}^{K}\lambda_2\boldsymbol{\nu}_k\|\mathbf{\Gamma}^k\|_1$$

- $\omega_j = \frac{1}{|\tilde{B}_{\ell_j,j}|}$
- $\nu_{k,j} = \frac{1}{|\tilde{B}_{k,j}-\tilde{B}_{\ell_j,j}|}$
- $\tilde{B}_{\ell_j,j} = \text{median}(\tilde{B}_{1,j}, \ldots, \tilde{B}_{K,j})$
- $\tilde{B}_k$ are some consistent initial estimate from individual-level models
    - Can be obtained via MLE, Ridge, or Lasso

Multi-VAR Networks



Outline

1 Heterogeneity

2 Multi-VAR

3 Subgrouping Multi-VAR

4 Simulation Results

5 Future Directions

Motivation

- In addition to shared and unique effects, subgroups of individuals may exhibit qualitative/quantitative similarities
  - Diagnostic status
- Identification of subgroup-level effects can be of substantive interest
  - Matching prevention/intervention efforts to subgroup characteristics
  - Early warning sign for onset of depressive episode (Whichers et al., 2016, 2019)
- Subgroups are also of interest for predictive goals
  - If subgroups are present, accurate model recovery will improve predictive accuracy

Subgroup Partition

- We can add an additional partition:

$$\mathbf{\Phi}^k = \mathbf{\Gamma}^0 + \mathbf{\Gamma}^s + \mathbf{\Gamma}^k$$

$$s = 1, \ldots, S, \;\; k = 1, \ldots, K$$

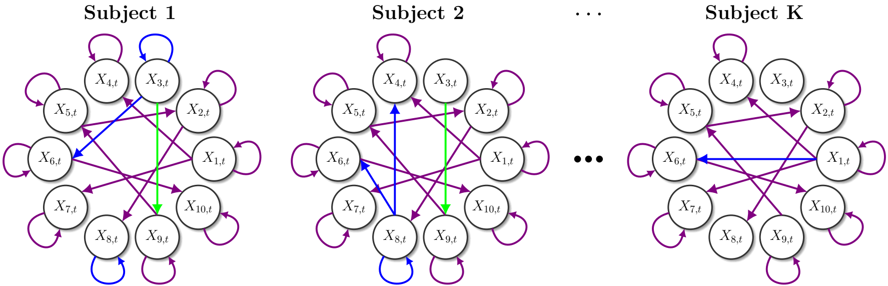
- $\mathbf{\Gamma}^s$ is a $d \times d$ matrix of subgroup effects for a given subgroup s
- Objective function now incorporates this further decomposition:

$$\frac{1}{N} \sum_{k=1}^{K} \|\mathbf{Y}^k - (\mathbf{\Gamma}^0 + \mathbf{\Gamma}^s + \mathbf{\Gamma}^k)\mathbf{Z}^k\|_2^2 + \lambda_1 \boldsymbol{\omega} \|\mathbf{\Gamma}^0\|_1 + \sum_{s=1}^{S} \lambda_2 \boldsymbol{\tau}_s \|\mathbf{\Gamma}^s\|_1 + \sum_{k=1}^{K} \lambda_3 \boldsymbol{\nu}_k \|\mathbf{\Gamma}^k\|_1$$

Walktrap

- Procedure for identifying subgroups and estimating subgroup-specific effects in a data-driven manner:
  1. Estimate transition matrices via Multi-VAR without subgrouping
  2. Create adjacency matrix where each element represents the number of shared effects between two people (presence and sign)
  3. Apply Walktrap algorithm to the adjacency matrix
  4. Estimate transition matrices via Multi-VAR using derived subgroups
- Walktrap is a random walk approach that merges communities in a bottom-up fashion using Ward's clustering (Pons & Lapaty, 2006; Ward, 1963)
- Performs well in other methodological frameworks characterized by heterogeneous time series (e.g., Gates et al., 2017; Lane et al., 2019; Park et al., 2022)

Subgroup Networks



Outline

1. Heterogeneity
2. Multi-VAR
3. Subgrouping Multi-VAR
4. Simulation Results
5. Future Directions

Simulation Design

- **Number of timepoints**: 50, 100
- **Number of individuals**: 10, 30
- **Number of subgroups**: 2, 3

- For 3 subgroup condition: 20%, 20%, 60%
- 10 variables for each condition
- 50 iterations for each condition

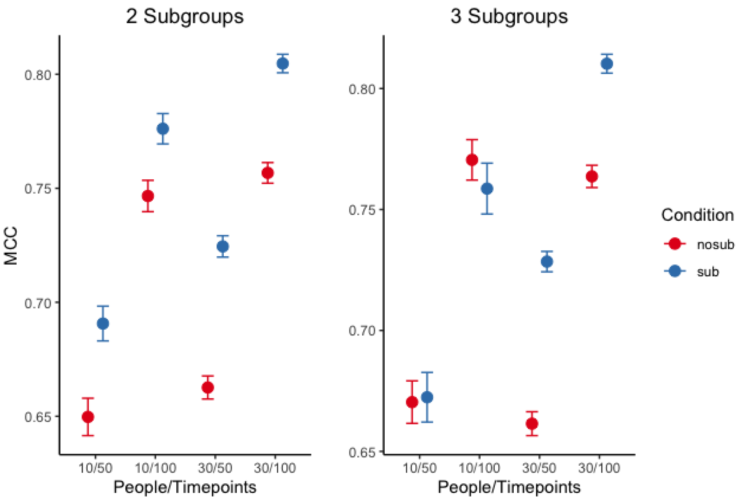
- Data *with* subgroups present simulated
- Adaptive Multi-VAR with *and* without subgrouping fit to data
- Initial estimates for adaptive weights obtained using Lasso

Simulation Design

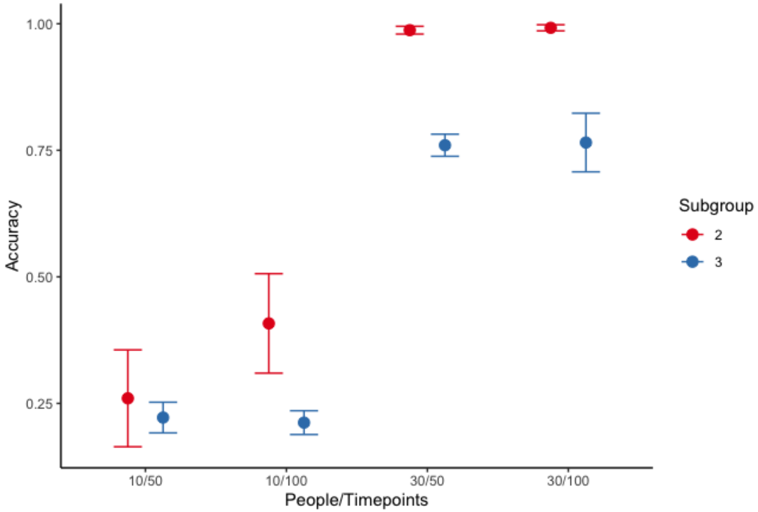
- **Group Effects**: All autoregressive paths
- **Subgroup Effects**: 5% of possible paths
- **Individual Effects**: 5% of possible paths
- All effects drawn from $\mathcal{U}(0, 1)$

$$\underbrace{\begin{bmatrix} 0.1 & 0 & 0 \\ 0.3 & 0.4 & 0 \\ 0 & 0.2 & 0.6 \end{bmatrix}}_{\text{Subject 1}} \underbrace{\begin{bmatrix} 0.2 & 0 & 0.1 \\ 0 & 0.8 & 0 \\ 0 & 0.4 & 0.1 \end{bmatrix}}_{\text{Subject 2}} \underbrace{\begin{bmatrix} 0.5 & 0.6 & 0 \\ 0 & 0.1 & 0.2 \\ 0 & 0 & 0.3 \end{bmatrix}}_{\text{Subject 3}}$$

Confirmatory Results



Subgroup Recovery



Outline

1 Heterogeneity

2 Multi-VAR

3 Subgrouping Multi-VAR

4 Simulation Results

5 Future Directions

Future Directions

- Possible remedies to poor subgroup recovery in small-sample and unbalanced scenarios
- Multi-VAR with subgrouping applied to real data
  - E.g., Diagnostic data
- Exploration of alternative algorithms for data-driven subgrouping
  - E.g., Fused Lasso

Thank you!