Subgrouping with Chain Graphical VAR

Jonathan J. Park
The Pennsylvania State University
Department of Human Development and Family Studies
JPark@psu.edu

February 24, 2021

Overview

- Idio-thetic Methods
- The VAR
- The Alternating Least Squares (ALS) VAR
- The Graphical VAR (gVAR)
- The Chain Graphical VAR (cgVAR)
- Subgrouping with cgVAR
- Demonstration

Sacha Epskamp
@SachaEpskamp

Following

Replying to @EikoFried

I mainly dislike representing contemporaneous effects as directed. While these are identified (lagged variables act as exogenous predictors identifying even cycles), exploratory estimation I think may often pick up the wrong direction. GIMME with graphical VAR would be cool!

3:29 AM - 14 Feb 2018

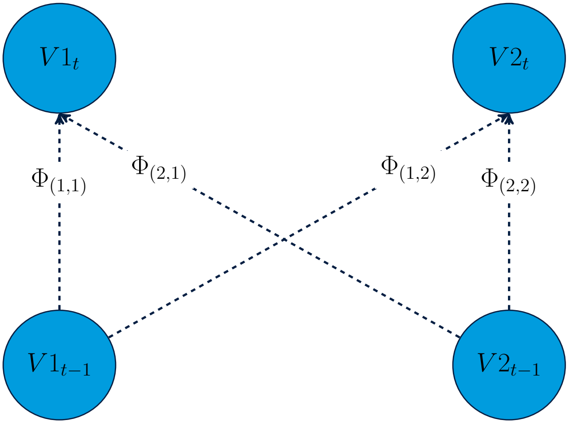
2 Likes

1 2

Idio-thetic Methods

- A class of methods that pool intraindividual information to derive nomothetic inference or vice versa
  - E.g., the multilevel VAR, the multi-VAR, GIMME and S-GIMME, the ALS VAR
  - All differ in allowances for more or less individual variability as well as in estimation
- Current challenge: undiagnosed heterogeneity can bias nomothetic generalizations from idio-thetic approaches
  - E.g., Distinct Profiles: *MDD* vs *Controls*
  - E.g., Sub-Profiles: MDD_1 vs MDD_2

Standard VAR Model



Standard VAR

$$\eta_t = c + \Phi\eta_{t-1} + \zeta_t$$

- $\eta_t = p$ variate vector of scores at time, t
- $c = p$ variate vector of constants
- $\Phi = p \times p$ dimensional matrix of lagged regression coefficients
- $\eta_{t-1} = p$ variate vector of scores at a given lag
- $\zeta_t = p$ variate vector of residuals

$$\zeta \sim N(0, \Psi)$$

Alternating Least Squares VAR

Bulteel et al., 2016

$$\eta_{it} = \sum_{k=1}^{K} p_{ik}(\mu_k + \Phi_k(\eta_{it-1} - \mu_i) + \zeta_{kt})$$

where

- $\eta_{it} = p$ variate vector of scores at time, t, for the i^{th} subject
- $p_{ik} =$ the $I \times K$ cluster-specific partition matrix
- $\mu_k = p$ variate vector of constants for the k^{th} subgroup
- $\Phi_k = p \times p$ dimensional matrix of lagged regression coefficients for the k^{th} subgroup
- $\zeta_{kt} = p$ variate vector of residuals

ALS VAR

Estimation

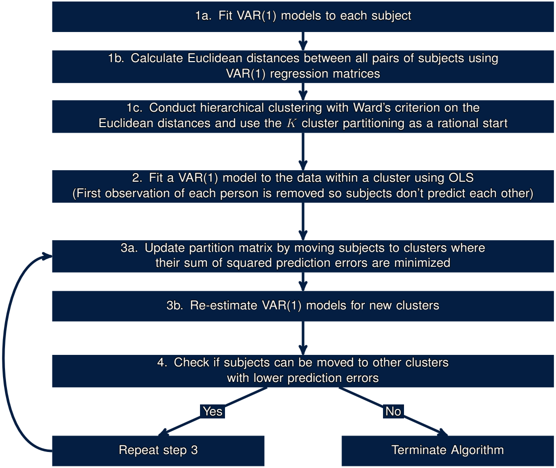
$$L_K = \sum_{i=1}^{I} \sum_{t=2}^{T} (\eta_{it} - \hat{\eta}_{it})^2$$

where

- $L_K =$ the sum of squared prediction errors
- $\hat{\eta}_{it} =$ the $p \times 1$ vector of predicted scores for the i^{th} subject at time, t

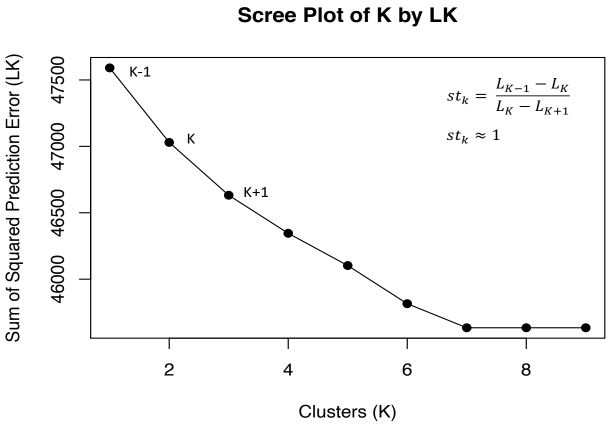
ALS VAR

Estimation



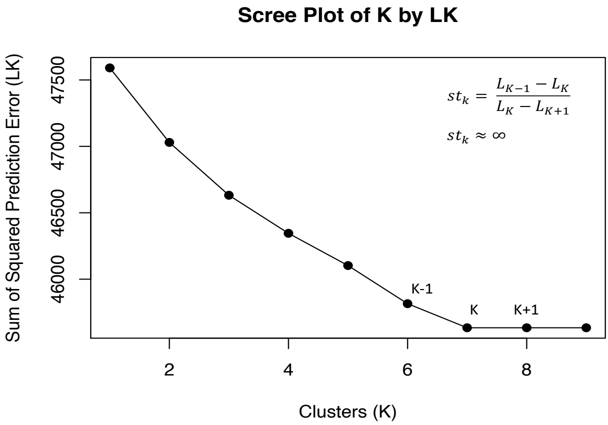
ALS VAR

Model Selection by scree ratio



ALS VAR

Model Selection by scree ratio

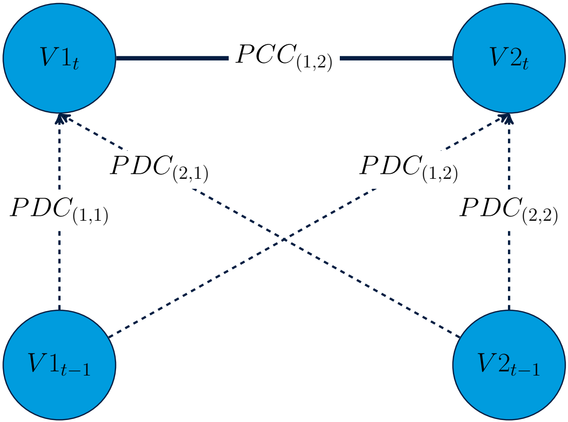


ALS VAR

Recap

- Derives K-cluster solution for VAR models
- "Forces" common structure to all subjects within a cluster
- Best models minimize prediction error while attempting to preserve parsimony

Graphical VAR



Graphical VAR

- Graphical VAR builds upon the VAR and estimates a network of contemporaneous partial correlations using the inverse residual covariance matrix
  - i.e., $\mathsf{K} = \Psi^{-1} = cov[\zeta_t, \zeta_t^t]^{-1}$
- Contemporaneous effects interpreted as X $\leftrightarrow$ Y conditioned upon all other pairwise associations
- Also generates a network of standardized, lagged regression coefficients

VAR → gVAR - The Partial Contemporaneous Network

The inverse of the residual covariance matrix (i.e., $K = \Psi^{-1}$) can be transformed into a network of partial contemporaneous correlations (PCCs) using the following:

$$PCC(X_{i,t}, X_{j,t}) = -\frac{K_{ij}}{\sqrt{K_{ii}K_{jj}}}$$

Where:

- K_{ij} = the element of K at coordinates (i, j)
- K_{ii} = the diagonal element of K associated with item, i
- K_{jj} = the diagonal element of K associated with item, j

VAR → gVAR - The Partial Directed Network

The lagged relationships of the VAR can also be standardized using information from K to form a network of partial directed correlations (PDCs) using the following formula:

$$PDC(X_{i,t}, X_{j,t-1}) = \frac{\Phi_{ij}}{\sqrt{\Psi_{ii} K_{jj} + \Phi_{ij}^2}}$$

Where:

- Φ_{ij} = the regression coefficient of i on j
- Ψ_{ii} = the residual variance of the outcome at time, t
- K_{jj} = the diagonal element of K associated with item, j

The Chain Graphical VAR

Epskamp et al., 2018

$$\eta_{i,t} = \mu_i + \Phi_i(\eta_{i,t-1} - \mu_i) + \zeta_{(i,t)}$$

where

- $\eta_{i,t}$ = p variate vector of scores at time, t, for subject i
- μ_i = the person-specific mean vector for subject, i
- $\Phi_i = p \times p$ dimensional matrix of lagged regression coefficients for subject i
- $\zeta_{i,t}$ = p variate residual vector at time, t for subject i

$$\zeta_{i,T} \sim N(0, \Psi_i)$$
$$\mathrm{K}_i^{(\Psi)} = \Psi_i^{-1}$$

Assumptions

Assuming a subject picked at random, and data are grand-mean centered, we expect the following (Epskamp et al., 2018):

$$\mathbb{E}(\mu_I) = 0$$
$$\mathbb{E}(\Phi_I) = \Phi_*$$
$$\mathbb{E}(K_I^{(\Psi)}) = K_*^{(\Psi)}$$

where

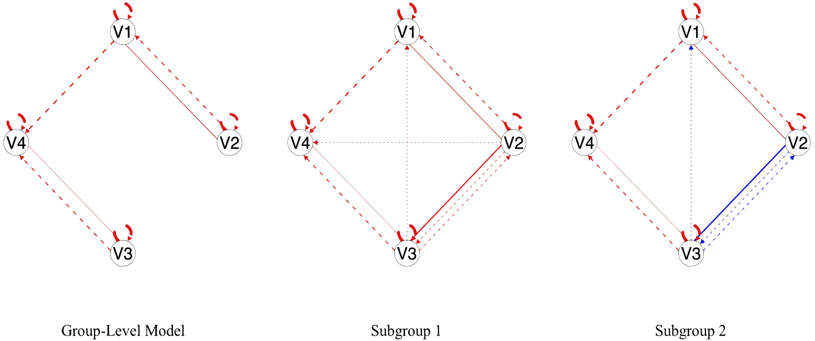
- $\Phi_* =$ average $p \times p$ dimensional lagged effects matrix
- $K_*^{(\Psi)} =$ average $p \times p$ dimensional precision matrix

e.g., $(\Phi_i - \Phi_*)$ would be the 'random effects'

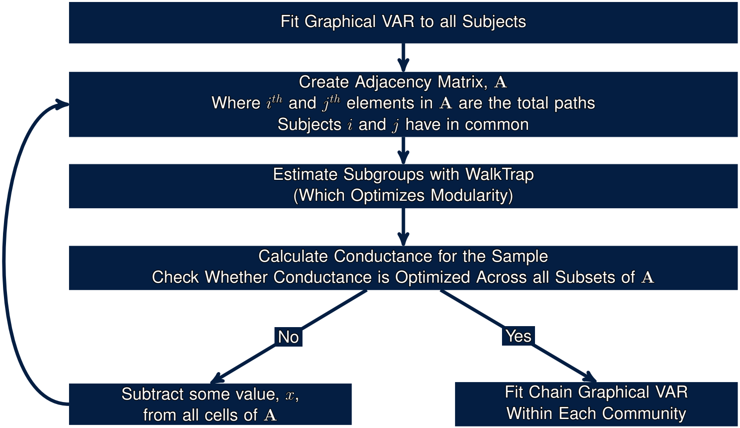
Recap

- A graphical VAR model can be fit to the chained time-series of multiple subjects; the chained graphical VAR
- Resulting "average" lagged and contemporaneous networks are thought of as common structures but are not imposed on subject-level networks
- Strong assumption of homogeneity to fit a chain gVAR
- Output contains parameterized networks at the group- and individual-level

Why Cluster?


V1
V2
V3
V4
Group-Level Model
V1
V2
V3
V4
Subgroup 1
V1
V2
V3
V4
Subgroup 2

Our Approach



Optimization of A

Why?

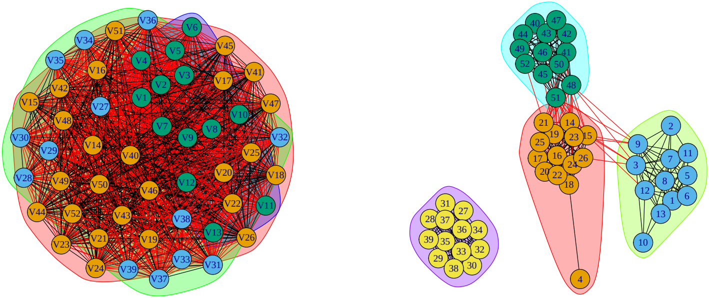
- Communities should be more densely connected to same-community members than they are to members of other communities
- S-GIMME, by default, subtracts the minimum value from all cells to induce sparsity
- This makes sense as we would expect the minimum value to exist in the space between communities

The Adjacency Matrix

0	15	13	18	4	6	5	4
15	0	11	12	3	4	4	3
13	11	0	12	2	2	2	2
18	12	12	0	3	4	2	4
4	3	2	3	0	9	6	9
6	4	2	4	9	0	8	10
5	4	2	2	6	8	0	6
4	3	2	4	9	10	6	0

Comparison of Adjacency Matrix Optimization

Simulated example - Minimum out versus Conductance



Recap on scgVAR

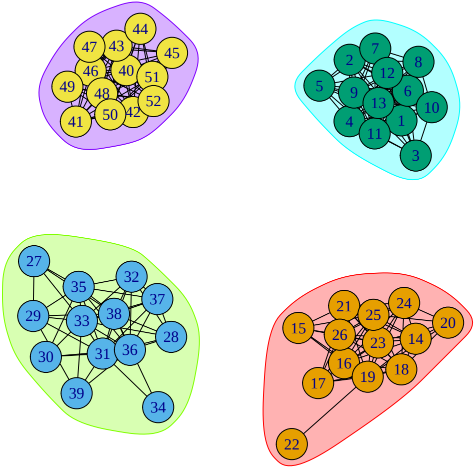
- scgVAR identifies homogeneous subgroups by optimizing the conductance of the person-by-person graph
- Fits a chain graphical VAR to the chained time-series of all individuals within each subgroup
- Provides 1 group-level network, K subgroup-level networks, and N person-specific networks

Simulated Illustration

- $N = 52$
- Network Size = 10-nodes
- 4-Simulated Subgroups
- $N_{reps} = 30$
- $T = 500$
- 10 autoregressions and 9 cross-regressions
  - 8 subgroup-specific
  - 1 shared between groups 1 and 2
  - 1 shared between groups 3 and 4

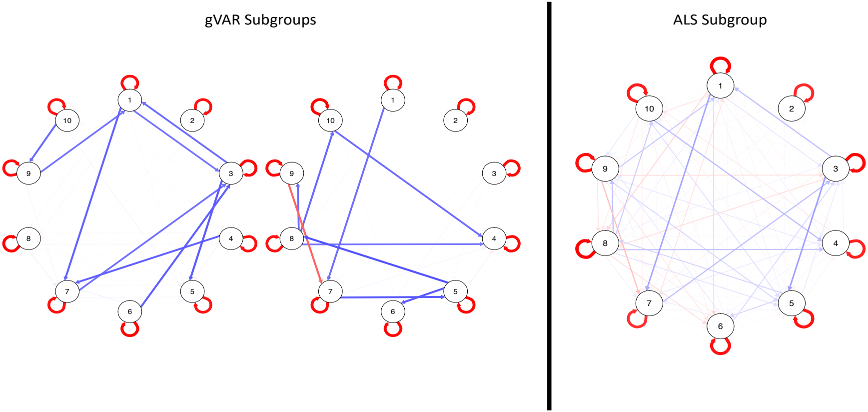
Simulated Illustration

Recovered Subgroups scgVAR



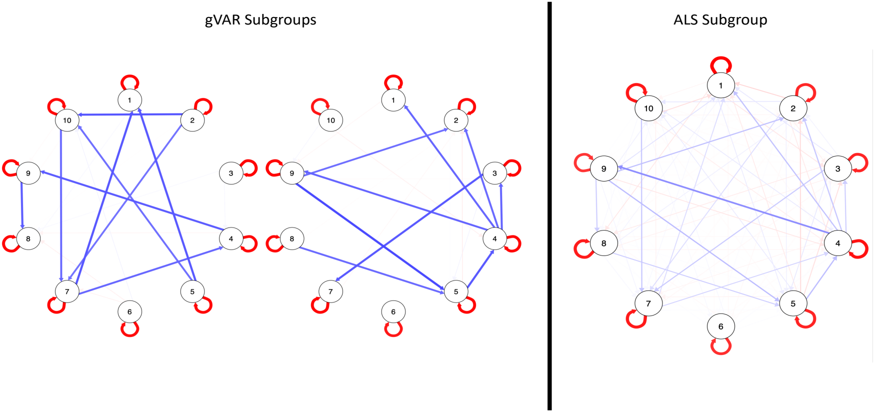
Simulated Illustration

scgVAR subgroups 1 and 2 == ALS VAR subgroup 1



Simulated Illustration

scgVAR subgroups 3 and 4 == ALS VAR subgroup 2

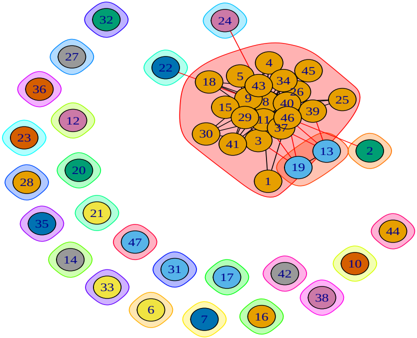


The MOOVD Study

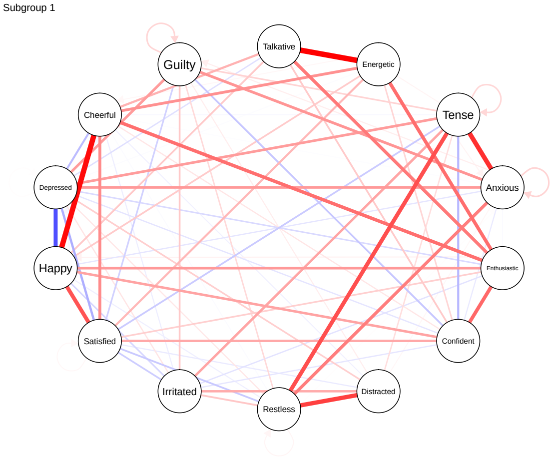
de Vos et al., 2017

- $N = 47$ participants
  - 24 participants had Major Depressive Disorder (MDD)
  - 23 participants were pair-matched controls
- $\bar{T} = 83.2$; $SD = 7.4$; measurements 3-times a day for 30-days
- Assessed on 14-affect items (7 positive; 7 negative)
- scgVAR settings:
  - $\gamma = 0.00$ model selection with BIC
  - $n\lambda = 10$; search 10×10 grid of possible λ_1 and λ_2 values
- ALS VAR settings:
  - $K_{max} = 47$; search all possible cluster combinations
  - Cluster solution which maximized st_k selected as optimal model

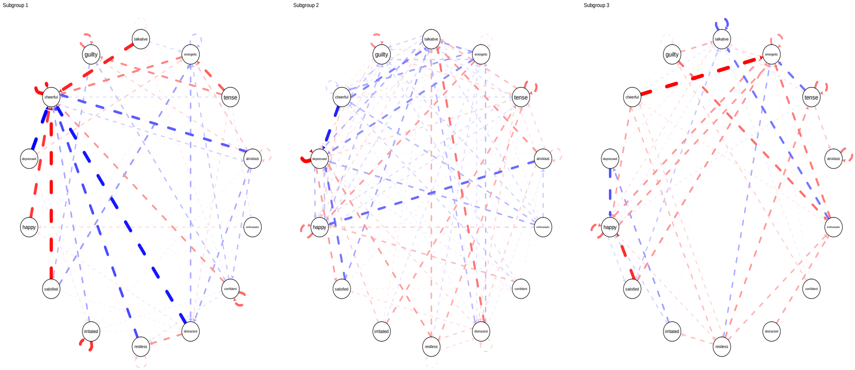
Demonstration



Demonstration



Demonstration



Concluding Remarks

- Idio-thetic methods allow for nomothetic inferences to be made by pooling intraindividual information
- We introduce Subgrouping with Chain Graphical VAR models as one way of making idio-thetic inference
- Will be coming to the graphicalVAR package soon :D
  - Can e-mail me for current prototype
- Feedback and inquiries can be sent to:
  - JPark@psu.edu
  - JonathanPark.dev

Thank you!