Time-Varying Mixed Graphical Models

Jonas Haslbeck and Lourens Waldorp

Psychological Methods
University of Amsterdam, the Netherlands

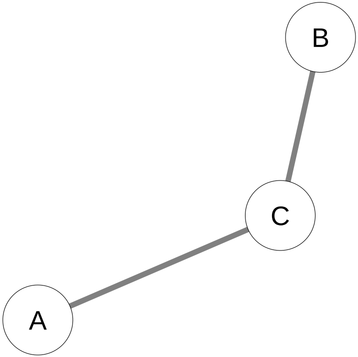
COMPSTAT 2016, August 26th

What are Graphical Models?

$$X_A \perp\!\!\!\perp X_B | X_C$$

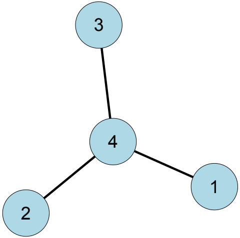
$$X_A \not\perp\!\!\!\perp X_C | X_B \quad \Longleftrightarrow$$

$$X_C \not\perp\!\!\!\perp X_B | X_A$$



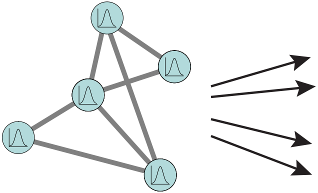
Example: Gaussian Graphical Model

$$\boldsymbol{\Sigma}^{-1} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 3.45 & 0 & 0 & 3.18 \\ 0 & 2.14 & 0 & 0.82 \\ 0 & 0 & 3.21 & 1.05 \\ 3.18 & 0.82 & 1.05 & 8.77 \end{pmatrix} \end{array} \Longleftrightarrow$$



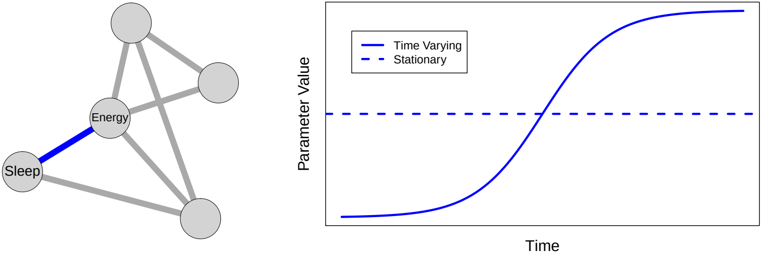
$$P(X_1, \ldots, X_p) = \frac{1}{\sqrt{(2\pi)^p |\Sigma|}} \exp\left\{ -\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right\}$$

Time-**in**variant Graphical Model

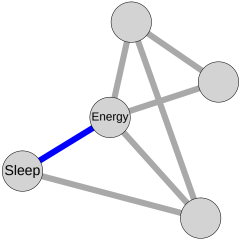


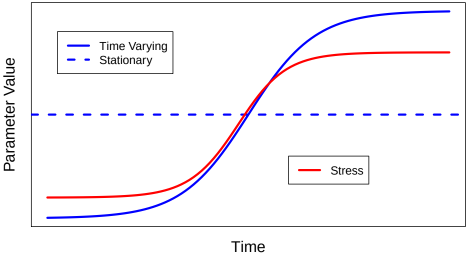
Time	X_1	X_2	X_3	X_4	X_5
1	3.45	1	0.98	3	1
2	1.11	3	0.82	3	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T-1$	0.12	2	0.71	2	2
T	-0.78	1	0.18	1	1

Parameters may change over time!

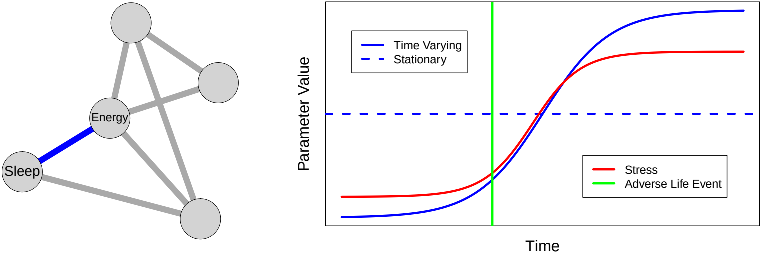


Parameters may change over time!





Parameters may change over time!

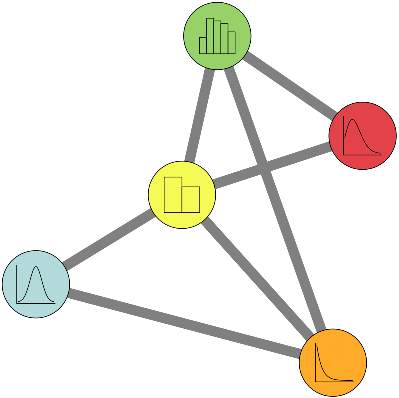


Introducing time-varying Mixed Graphical Models

1) Stationary Mixed Graphical Models

2) Extension to the time-varying case

Mixed Exponential Graphical Model



Construction of Mixed Exponential Graphical Model

Conditional univariate members of the exponential family

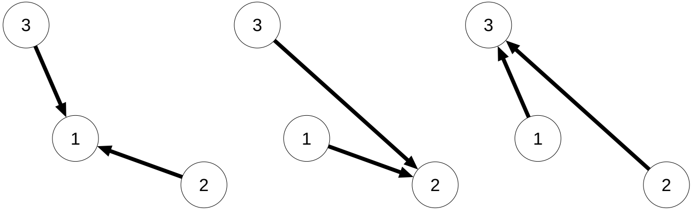
$$P(X_s|X_{\setminus s}) = \exp\left\{E_s(X_{\setminus s})\phi_s(X_s) + C_s(X_s) - \Phi(X_{\setminus s})\right\},$$

factorize to a global multivariate distribution which factors according the graph defined by the node-neighborhoods if and only if $E_s(X_{\setminus s})$ has the form:

$$\theta_s + \sum_{t \in N(s)} \theta_{st}\phi_t(X_t) + \ldots + \sum_{t_2,\ldots,t_k \in N(s)} \theta_{t_2,\ldots,t_k} \prod_{j=2}^{k} \phi_{t_j}(X_{t_j}),$$

(Yang et al., 2014)

Neighborhood Regression Method



(Meinshausen & Buehlmann, 2006)

Estimation Algorithm

For each node s :

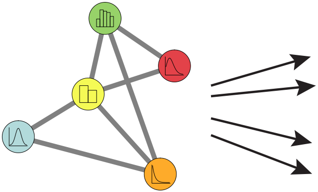
1. Regress $X_{\setminus s}$ on X_s
   - $\min_{(\theta_0,\theta)\in\mathbb{R}^p} \left[\frac{1}{N}\sum_{i=1}^{N}(y_i - \theta_0 - X_{\setminus s;i}^T\theta)^2 + \lambda_n ||\theta||_1\right]$
   - Select λ_n using EBIC
2. Threshold Parameter Estimates
   - $\tau_n \asymp \sqrt{d}||\theta||_2\sqrt{\frac{\log p}{n}}$

Combine Estimates from both regressions

- AND-rule: Edge present if both parameters are nonzero
- OR-rule: Edge present if at least one parameter is nonzero

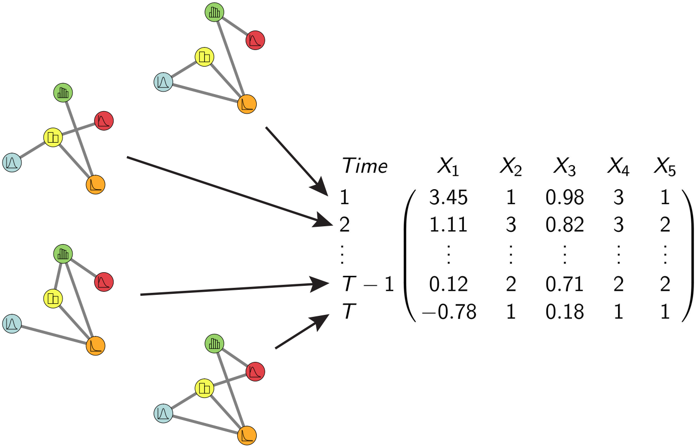
(Loh & Wainwright, 2013)

Recab: Time-**in**variant mixed Graphical Model

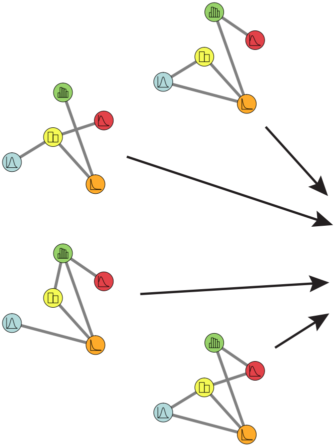


Time	X_1	X_2	X_3	X_4	X_5
1	3.45	1	0.98	3	1
2	1.11	3	0.82	3	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T-1$	0.12	2	0.71	2	2
T	-0.78	1	0.18	1	1

Time-varying mixed Graphical Model



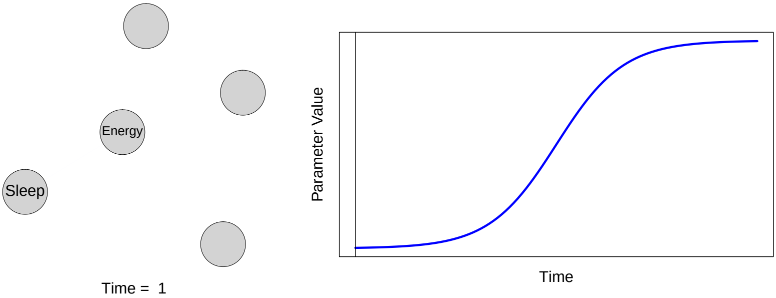
Time-varying mixed Graphical Model



Time	X_1	X_2	X_3	X_4	X_5
1	3.45	1	0.98	3	1
2	1.11	3	0.82	3	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T-1$	0.12	2	0.71	2	2
T	-0.78	1	0.18	1	1

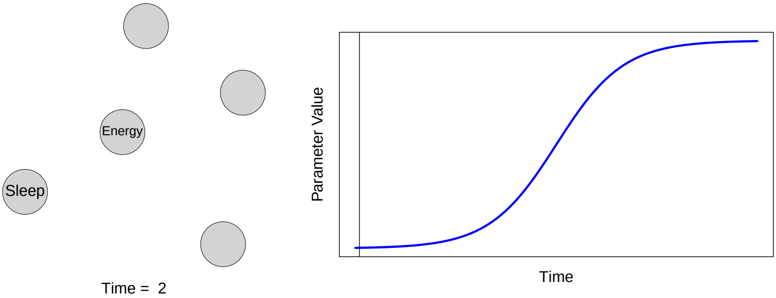
But: we have the scaling $\tau_n \asymp \sqrt{d}||\theta||_2\sqrt{\frac{\log p}{n}}$

Local Stationarity



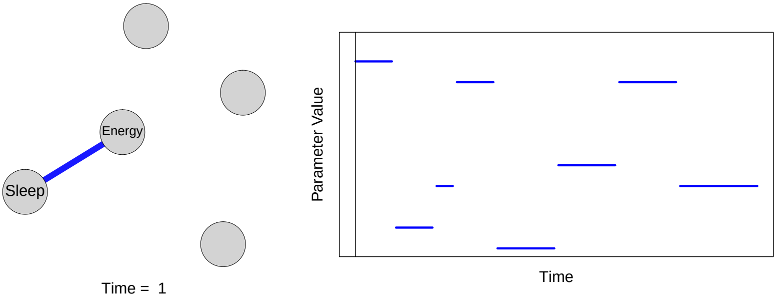
Assumption: Edge parameters are a smooth function of time

Local Stationarity



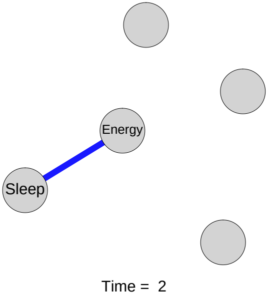
Assumption: Edge parameters are a smooth function of time

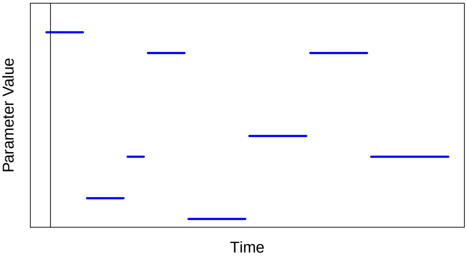
Local Stationarity Violated!



Edge parameters are *not* a smooth function of time!

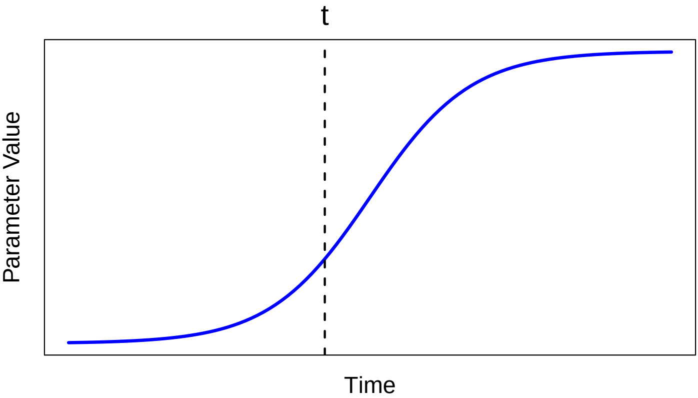
Local Stationarity Violated!



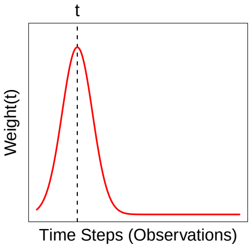


Edge parameters are *not* a smooth function of time!

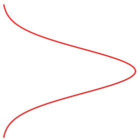
Again: Local Stationarity



Time-varying Graphs via Weighted Regression



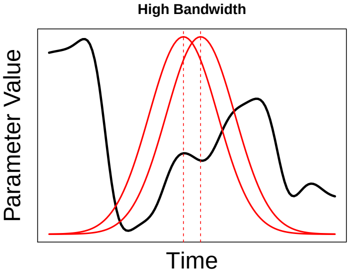
Time	X_1	...	X_5	*Weight*(t)
1	0.92	...	−1.47	0
2	0.78	...	−0.48	0.13
3	0.07	...	0.42	0.60
4	−1.99	...	1.36	1.00
5	0.62	...	−0.10	0.60
6	−0.06	...	0.39	0.13
7	−0.21	...	0.99	0
⋮	⋮	⋮	⋮	⋮
T	−0.16	...	0.18	0

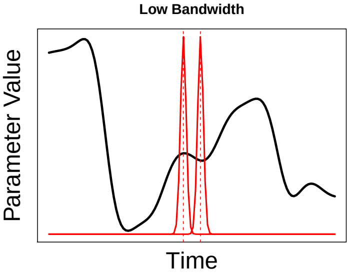


Weighted cost function:

$$\min_{(\theta_0,\theta)\in\mathbb{R}^p}\left[\frac{1}{N}\sum_{i=1}^{N} w_{i;t}(y_{i;t}-\theta_{0;t}-X_{\backslash s;i}^{T}\theta_t)^2+\lambda_n||\theta_t||_1\right]$$

What is the right bandwidth?

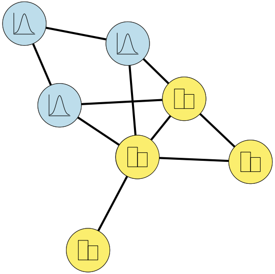




More Information for estimation   Higher sensitivity to change

Scaling: $\tau_n \asymp \sqrt{d}||\theta||_2\sqrt{\frac{\log p}{n}}$

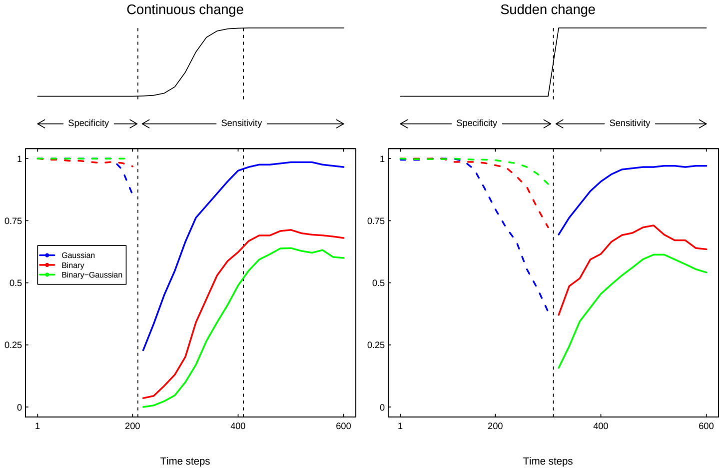
Determine Performance via Simulation



- Binary-Gaussian Graphical Model
- 20 Nodes
- Always 19 edges present
- Of these are always 6 changing
- Type of change: smooth vs. sudden

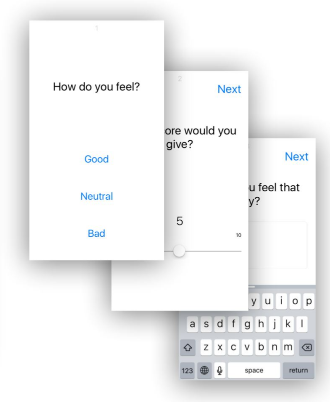
Simulation: Results

N = 600
Nt = 205.9
Nt/p = 10.3



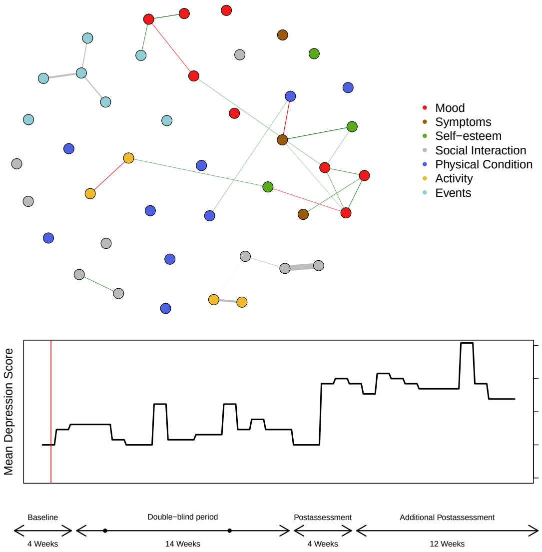
$$\text{bandwidth} = 0.8/N^{1/3} \approx 0.095$$

Application to Event Sampling Data



- Time series of 1 person
- 43 variables
- Up to 10 measurements a day
- For 36 weeks

Time-varying Graph of Psychopathology



Time-varying Mixed Graphical Models

Summary

- Estimation of time-varying mixed Graphical Models
- Assumption of local stationarity
- Works in realistic situations
- Available in 'mgm' R-package on CRAN

Contact

- Email: jonashaslbeck@gmail.com
- Website: http://jmbh.github.io

Mixed Graphical Model: Conditional Distribution

Conditional univariate members of the exponential family

$$P(X_s|X_{\setminus s}) = \exp\left\{E_s(X_{\setminus s})\phi_s(X_s) + C_s(X_s) - \Phi(X_{\setminus s})\right\},$$

factorize to a global multivariate distribution which factors according the graph defined by the node-neighborhoods if and only if $E_s(X_{\setminus s})$ has the form:

$$\theta_s + \sum_{t \in N(s)} \theta_{st}\phi_t(X_t) + \ldots + \sum_{t_2,\ldots,t_k \in N(s)} \theta_{t_2,\ldots,t_k} \prod_{j=2}^{k} \phi_{t_j}(X_{t_j}),$$

Mixed Graphical Model: Joint Distribution

The joint distribution has the form

$$P(X;\theta) = \exp\{\sum_{s\in V}\theta_s\phi_s(X_s) + \sum_{s\in V}\sum_{t\in N(s)}\theta_{st}\phi_s(X_s)\phi_t(X_t) +$$

$$\cdots + \sum_{t_1,\ldots,t_k\in\mathcal{C}}\theta_{t_1,\ldots,t_k}\prod_{j=1}^{k}\phi_{t_j}(X_{t_j}) + \sum_{s\in V}C_s(X_s) - \Phi(\theta)\}$$

Example Mixed Graphical Model: Ising-Gaussian

$$P(Y,Z) \propto \exp\Big\{ \sum_{s\in V_Y} \frac{\theta_s^y}{\sigma_s} Y_s + \sum_{r\in V_Z} \theta_r^z Z_r + \sum_{(s,t)\in E_Y} \frac{\theta_{st}^{yy}}{\sigma_s\sigma_t} Y_s Y_t + \sum_{(r,q)\in E_Z} \theta_{rq}^{zz} Z_r Z_q + \sum_{(s,r)\in E_{YZ}} \frac{\theta_{sr}^{yz}}{\sigma_s} Y_s Z_r - \sum_{s\in V_Y} \frac{Y_s^2}{2\sigma_s^2} \Big\}$$

If X_s Bernoulli, the node-conditional has the form:

$$P(X_s|X_{\setminus s}) \propto \exp\Big\{ \theta_r^z Z_r + \sum_{q\in N(r)_Z} \theta_{rq}^{zz} Z_r Z_q + \sum_{t\in N(r)_Y} \frac{\theta_{rt}^{yz}}{\sigma_t} Z_r Y_t \Big\}$$

If X_s Gaussian, the node-conditional has the form:

$$P(X_s|X_{\setminus s}) \propto \exp\Big\{ \frac{\theta_s^y}{\sigma_s} Y_s + \sum_{t\in N(s)_Y} \frac{\theta_{st}^{yy}}{\sigma_s\sigma_t} Y_s Y_t + \sum_{r\in N(s)_Z} \frac{\theta_{sr}^{yz}}{\sigma_s} Y_s Z_r - \frac{Y_s^2}{2\sigma_s^2} \Big\}$$