Structure Estimation for Mixed Graphical Models

Jonas Haslbeck [1] Lourens Waldorp [2]

[1]Department of Methodology & Statistics, Utrecht University

[2]Department of Psychological Methods, University of Amsterdam

September 1, 2015

Overview

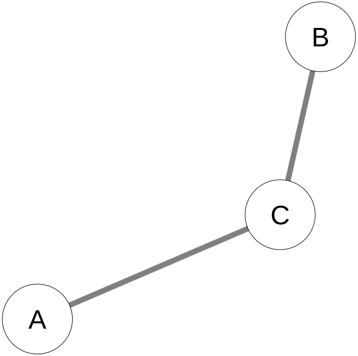
1. Background and current limitations
2. Theory: Modeling and estimating mixed graphical models
3. Performance benchmarks
4. Application to Autism dataset
5. Implementation: R-package

Markov random fields

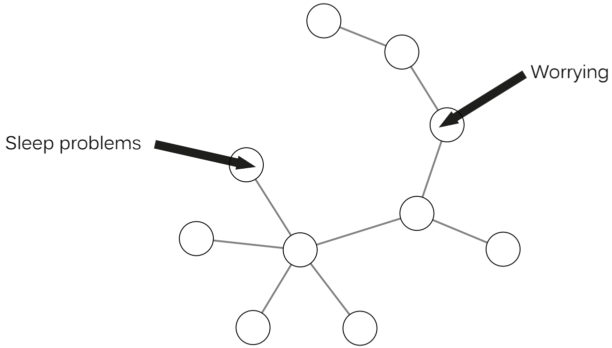
$$X_A \perp\!\!\!\perp X_B | X_C$$

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

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

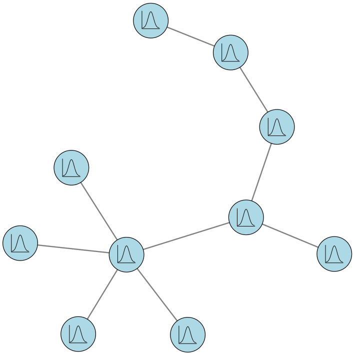


Why are Markov random fields interesting?

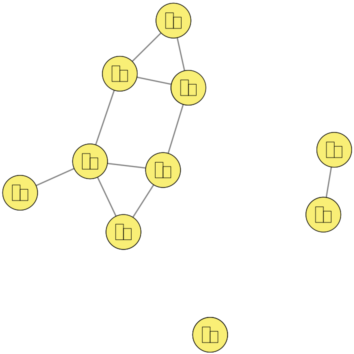


Depression Symptoms

Existing structure estimation methods

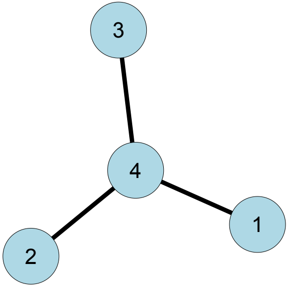


Gaussian



Binary (Ising model)

Structure estimation in the Gaussian case

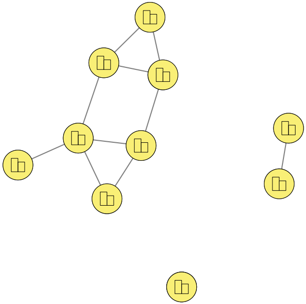


$$\Longleftrightarrow \quad \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 3.45 & 0 & 0 & 3.18 \\ 0 & 2.14 & 0 & 0.82 \\ 0 & 0 & 3.21 & 1.05 \\ 3.18 & 0.82 & 1.05 & 8.77 \end{pmatrix} \end{array}$$

Estimation:

- glasso (Friedman et al., 2008)
- Nodewise methods (Meinshausen & Bühlmann, 2006)

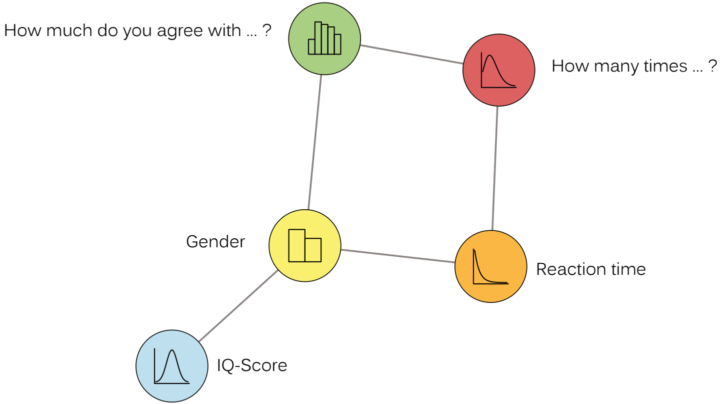
Structure estimation for the Ising model



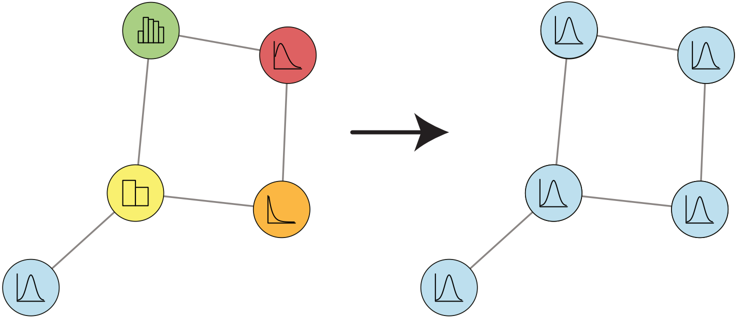
Estimation via nodewise methods:

- ℓ_1-regularized logistic regression (Ravikumar et al.,2010)
- eLasso (van Borkulo et al., 2014)

Mixed Markov random fields



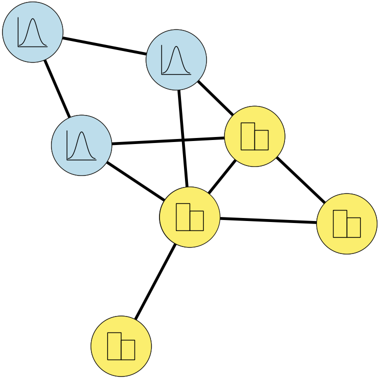
Gaussianizing variables



Two approaches:

- Copula-based (Dobra and Lenosti, 2001; Liu et al. 2012)
- Non-paranormal (Liu et al., 2009; Lafftery et al. 2012)

Conditional Gaussian

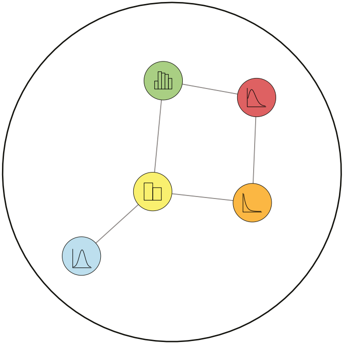


Multivariate gaussian conditioned on $2^{|Binary\ nodes|}$ configurations.

Improvements:

- Only pairwise interactions (Lee and Hastie, 2012)
- Only up to three-way interactions (Cheng et al., 2013)

How to estimate mixed Markov random fields in a principled way?



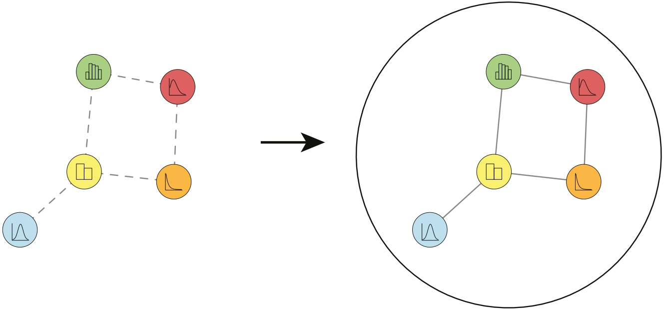
Modeling

+

$$\Gamma = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{array} \begin{array}{c} \begin{array}{ccccc} X_1 & X_2 & X_3 & X_4 & X_5 \end{array} \\ \begin{pmatrix} 4.04 & 1.05 & 0 & 3.18 & 0 \\ 1.05 & 1.44 & 2.75 & 0.82 & 0.79 \\ 0 & 2.75 & 5.52 & 1.05 & 4.22 \\ 3.18 & 0.82 & 1.05 & 4.50 & 3.88 \\ 0 & 0.79 & 4.22 & 3.88 & 3.18 \end{pmatrix} \end{array}$$

Estimation

Modeling mixed Markov random fields



Conditional univariate distributions → Joint distribution

(Yang et al., 2014)

Mixed exponential Markov random fields

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:

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

where $\theta_{s\cdot} := \{\theta_s, \theta_{st}, ..., \theta_{st_2...t_k}\}$ is a set of parameters and $N(s)$ is the set of neighbors of node s according to graph G.

Mixed exponential Markov random fields

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:

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

where $\theta_{s\cdot} := \{\theta_s, \theta_{st}, \ldots, \theta_{st_2\ldots t_k}\}$ is a set of parameters and $N(s)$ is the set of neighbors of node s according to graph G.

Mixed exponential Markov random fields

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:

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

where $\theta_{s\cdot} := \{\theta_s, \theta_{st}, \ldots, \theta_{st_2\ldots t_k}\}$ is a set of parameters and $N(s)$ is the set of neighbors of node s according to graph G.

Mixed exponential Markov random fields

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:

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

where $\theta_{s\cdot} := \{\theta_s, \theta_{st}, \dots, \theta_{st_2\dots t_k}\}$ is a set of parameters and $N(s)$ is the set of neighbors of node s according to graph G.

Mixed exponential Markov random fields

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:

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

where $\theta_{s\cdot} := \{\theta_s, \theta_{st}, \dots, \theta_{st_2\dots t_k}\}$ is a set of parameters and $N(s)$ is the set of neighbors of node s according to graph G.

Mixed exponential Markov random fields

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:

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

where $\theta_{s\cdot} := \{\theta_s, \theta_{st}, \ldots, \theta_{st_2\ldots t_k}\}$ is a set of parameters and $N(s)$ is the set of neighbors of node s according to graph G.

Mixed exponential Markov random fields

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:

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

where $\theta_{s\cdot} := \{\theta_s, \theta_{st}, ..., \theta_{st_2...t_k}\}$ is a set of parameters and $N(s)$ is the set of neighbors of node s according to graph G.

Mixed exponential Markov random fields

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:

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

where $\theta_{s\cdot} := \{\theta_s, \theta_{st}, ..., \theta_{st_2...t_k}\}$ is a set of parameters and $N(s)$ is the set of neighbors of node s according to graph G.

Mixed exponential Markov random fields: Joint

The corresponding joint distribution has the form:

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

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

where $\Phi(\theta)$ is the log-normalization constant.

Mixed exponential Markov random fields: Joint

The corresponding joint distribution has the form:

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

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

where $\Phi(\theta)$ is the log-normalization constant.

Mixed exponential Markov random fields: Joint

The corresponding joint distribution has the form:

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

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

where $\Phi(\theta)$ is the log-normalization constant.

Mixed exponential Markov random fields: Joint

The corresponding joint distribution has the form:

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

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

where $\Phi(\theta)$ is the log-normalization constant.

Mixed exponential Markov random fields: 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:

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

Mixed exponential Markov random fields: 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:

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

Mixed exponential Markov random fields: 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:

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

Mixed exponential Markov random fields: 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:

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

Mixed exponential Markov random fields: Ising-Gaussian

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

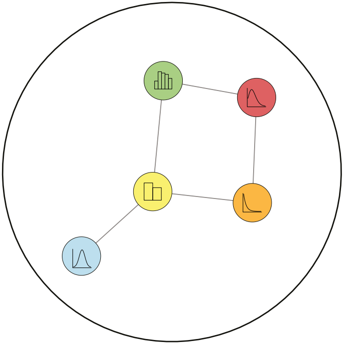
If X_s Bernoulli, the node-conditional has the form:

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

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

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

How to estimate mixed Markov random fields in a principled way?



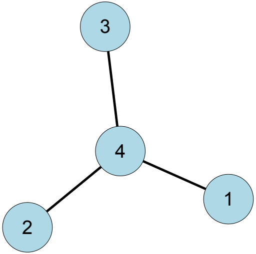
Modeling

+

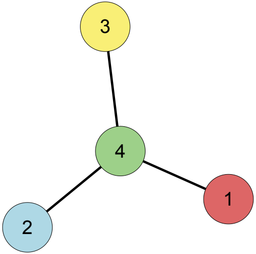
$$\Gamma = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{array} \begin{array}{c} \begin{array}{ccccc} X_1 & X_2 & X_3 & X_4 & X_5 \end{array} \\ \begin{pmatrix} 4.04 & 1.05 & 0 & 3.18 & 0 \\ 1.05 & 1.44 & 2.75 & 0.82 & 0.79 \\ 0 & 2.75 & 5.52 & 1.05 & 4.22 \\ 3.18 & 0.82 & 1.05 & 4.50 & 3.88 \\ 0 & 0.79 & 4.22 & 3.88 & 3.18 \end{pmatrix} \end{array}$$

Estimation

Inverse covariance matrices and graph structure

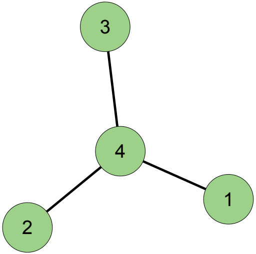


$$\Longleftrightarrow \quad \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 3.45 & 0 & 0 & 3.18 \\ 0 & 2.14 & 0 & 0.82 \\ 0 & 0 & 3.21 & 1.05 \\ 3.18 & 0.82 & 1.05 & 8.77 \end{pmatrix} \end{array}$$

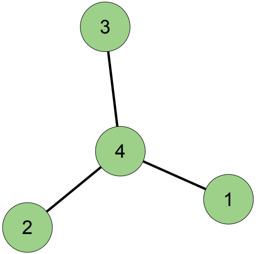


$$\not\Longleftrightarrow \quad \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 3.45 & 5.12 & 0 & 3.18 \\ 5.12 & 2.14 & 0.51 & 0.82 \\ 0 & 0.51 & 3.21 & 1.05 \\ 3.18 & 0.82 & 1.05 & 8.77 \end{pmatrix} \end{array}$$

Generalized covariance matrices



$$\not\Longleftrightarrow \quad \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 3.45 & 5.12 & 0 & 3.18 \\ 5.12 & 2.14 & 0.51 & 0.82 \\ 0 & 0.51 & 3.21 & 1.05 \\ 3.18 & 0.82 & 1.05 & 8.77 \end{pmatrix} \end{array}$$



$$\Longleftrightarrow \quad \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_1X_2 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} X_1 & X_2 & X_3 & X_4 & X_1X_2 & \dots \end{array} \\ \begin{pmatrix} 3.45 & 0 & 0 & 3.18 & 4.98 & \dots \\ 0 & 2.14 & 0 & 0.82 & 1.15 & \dots \\ 0 & 0 & 3.21 & 1.05 & 4.48 & \dots \\ 3.18 & 0.82 & 1.05 & 8.77 & 4.37 & \dots \\ 4.98 & 1.15 & 4.48 & 4.37 & 8.45 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array}$$

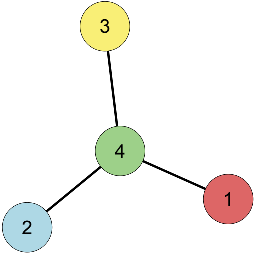
(Loh and Wainwright, 2013)

Generalized covariance matrices & nodewise regression

Corollary:
For any graph with maximal degree d, the inverse Γ of the covariance matrix over the node s and all its candidate neighborhoods up to size d is graph structured with respect to the $N(s)$. That is, $\Gamma(s,t) = 0$ whenever $t \notin N(s)$.

$$\begin{array}{c} \\ X_s \\ X_{t_1} \\ X_{t_2} \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccc} X_s & X_{t_1} & X_{t_2} & \dots \end{array} \\ \begin{pmatrix} 3.45 & 0 & 1.27 & \dots \\ 0 & 2.14 & 0 & \dots \\ 1.27 & 0 & 3.21 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array}$$

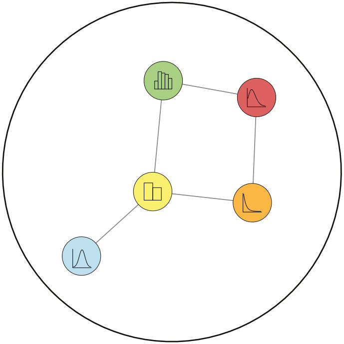
Generalized covariance matrices for mixed exponential Markov random fields



$\Longleftrightarrow$

	X_1	X_2	X_3	X_4	X_1X_2	$\dots$
X_1	3.45	0	0	3.18	4.98	$\dots$
X_2	0	2.14	0	0.82	1.15	$\dots$
X_3	0	0	3.21	1.05	4.48	$\dots$
X_4	3.18	0.82	1.05	8.77	4.37	$\dots$
X_1X_2	4.98	1.15	4.48	4.37	8.45	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

How to estimate mixed Markov random fields in a principled way?



Modeling

+

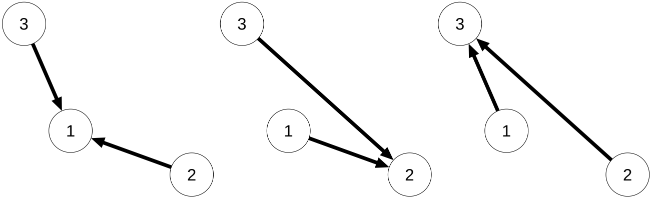
$$\Gamma = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{array} \begin{array}{c} \begin{array}{ccccc} X_1 & X_2 & X_3 & X_4 & X_5 \end{array} \\ \begin{pmatrix} 4.04 & 1.05 & 0 & 3.18 & 0 \\ 1.05 & 1.44 & 2.75 & 0.82 & 0.79 \\ 0 & 2.75 & 5.52 & 1.05 & 4.22 \\ 3.18 & 0.82 & 1.05 & 4.50 & 3.88 \\ 0 & 0.79 & 4.22 & 3.88 & 3.18 \end{pmatrix} \end{array}$$

Estimation

✓ ✓

Nodewise estimation algorithm

1. Regress all nodes $V_{\setminus s}$ on node V_s with a ℓ_1-penalty



2. Threshold parameters at $\tau_n = \sqrt{d}||\widehat{\beta}||_2\sqrt{\frac{\log p}{n}}$

3. Combine parameter estimates

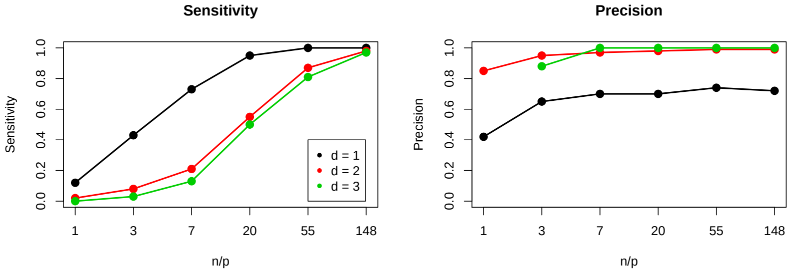
$$\widehat{\beta} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \end{array} \begin{array}{c} \begin{array}{ccc} X_1 & X_2 & X_3 \end{array} \\ \begin{pmatrix} NA & 0 & 4.78 \\ 0 & NA & 0.12 \\ 5.11 & 0 & NA \end{pmatrix} \end{array}$$

Simulation: Setup

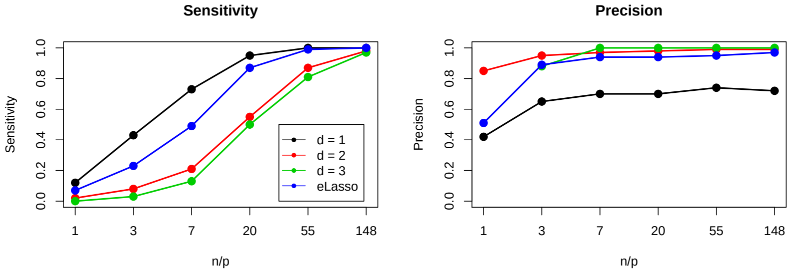
Varied factors:

1. Sparsity: $\{.1, .2, .3\}$
2. Ratio $\frac{n}{p}$: $\exp\{0, 1, 2, 3, 4, 5\} \approx \{1, 3, 7, 20, 55, 148\}$
3. Degree of augmented interactions d: $\{1, 2, 3\}$
4. Different (mixed) graphs
   - 4.1 Potts model with $m = \{2, 3, 4\}$
   - 4.2 Ising-Gaussian
   - 4.3 Ising-Exponential
   - 4.4 Ising-Poisson

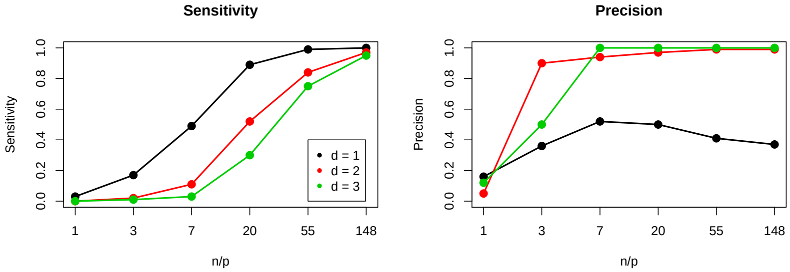
Results: Potts model (m=2) (Ising model)



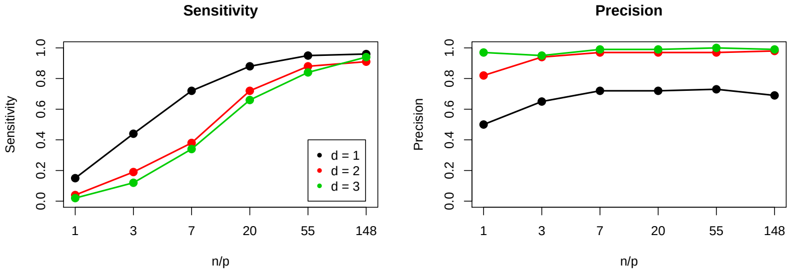
Results: Potts model (m=2) (Ising model)



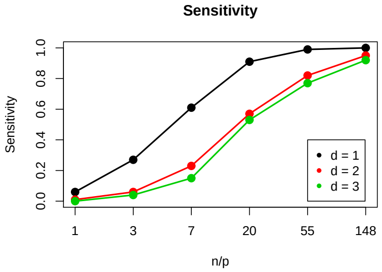
Results: Potts model (m=3)



Results: Ising-Gaussian



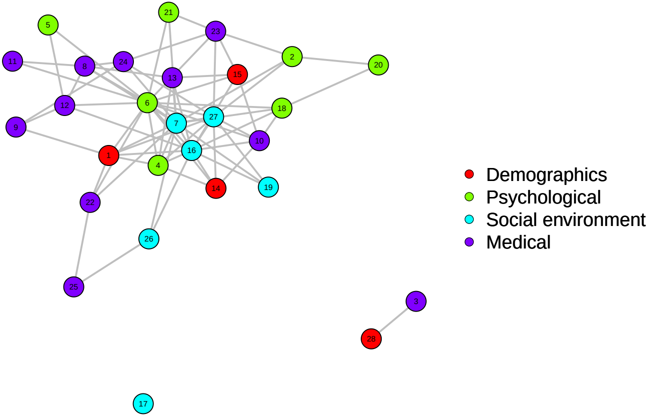
Results: Ising-Poisson



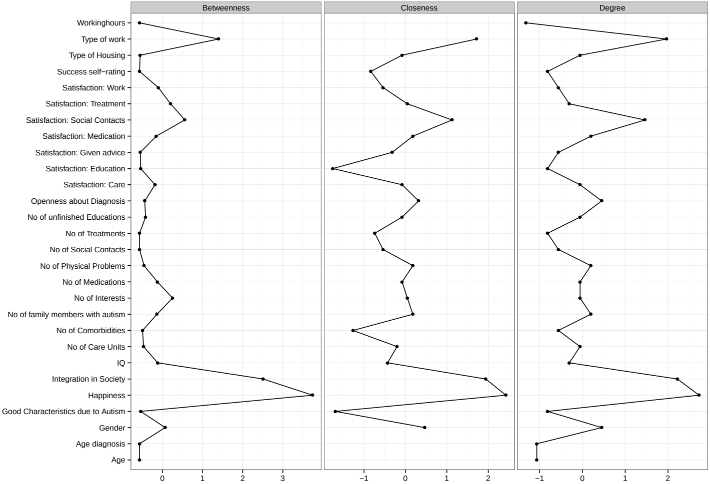
Exploring Autism-dataset

- 27 Variables describing the life of individuals diagnosed with Autism Spectrum Disorder (ASD) in the Netherlands (N=3521)
- Variables: Workinghours, Type of Work, Type of housing, Success, Satisfaction with Work, Satisfaction with treatment, Satisfaction with social contacts, Satisfaction with medication, Satisfaction with given advice, Satisfaction with education, Satisfaction with Care, Openness about diagnosis, Education, Number of social contacts, Physical problems, Medications, Interests, Family members with autism, Number of care units, IQ, Integration in Society, ...

Exploring Autism-dataset: Graph-visualization



Exploring Autism-dataset: Centrality-measures



R-package **mgm**

Install:

```
library(devtools)
install_github("jmbh/mgm")
library(mgm)
```

Fit a mixed Markov random field:

```
> round(head(data_mixed2),4)[1:3,]
        [,1] [,2]    [,3] [,4] [,5]
[1,]  0.6680    1 11.4234    2    2
[2,] -0.7114    1 22.9344    1    1
[3,]  1.2265    0 34.4966    3    1

type <- c("g", "p", "e", "c", "c")
levs <- c(1,1,1,3,2)

set.seed(5)
fit <- mgmfit(data = data_example, type = type, lev = levs,
              lambda.sel = "CV", folds = 10, gam = .25,
              d = 2, rule.reg = "AND", rule.cat = "OR")

?mgmfit
```

R-package **mgm**: Output

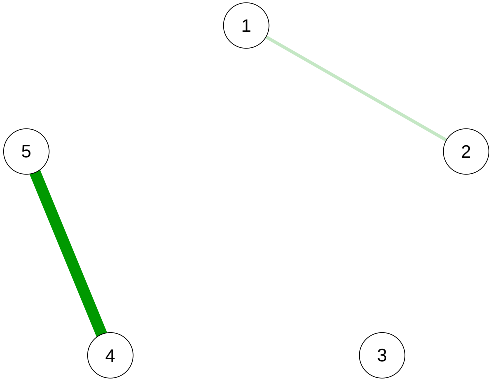
Output:

```
> fit
$adj
     [,1] [,2] [,3] [,4] [,5]
[1,]    0    1    0    0    0
[2,]    1    0    0    0    0
[3,]    0    0    0    0    0
[4,]    0    0    0    0    1
[5,]    0    0    0    1    0

$wadj
         [,1]     [,2] [,3]      [,4]      [,5]
[1,] 0.000000 0.177809    0 0.0000000 0.0000000
[2,] 0.177809 0.000000    0 0.0000000 0.0000000
[3,] 0.000000 0.000000    0 0.0000000 0.0000000
[4,] 0.000000 0.000000    0 0.0000000 0.7687597
[5,] 0.000000 0.000000    0 0.7687597 0.0000000

?mgmfit
```

R-package **mgm**: Visualize

Output:

```
library(qgraph)
qgraph(fit$wadj)
```



Summary

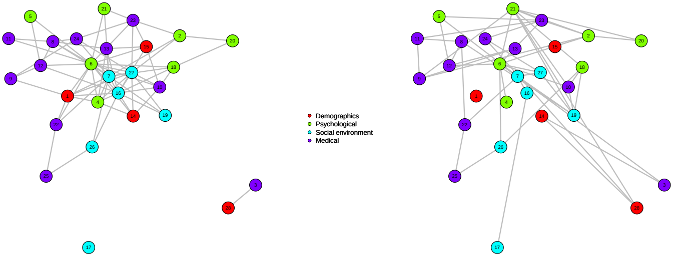
1. First "principled" method to estimate mixed Markov random fields
2. Performance measures: works in practical situations
3. R-package implementation: **mgm**

Contact:

jonashaslbeck@gmail.com
https://github.com/jmbh

Backup slides

Does proper modeling matter?



Mixed

All Gaussian

1: Gender
2: IQ
3: Age diagnosis
4: Openness about Diagnosis
5: Success self−rating
6: Happiness
7: Integration in Society
8: No of family members with autism
9: No of Comorbidities
10: No of Physical Problems
11: No of Treatments
12: No of Medications
13: No of Care Units
14: Type of Housing
15: No of unfinished Educations
16: Type of work
17: Workinghours
18: No of Interests
19: No of Social Contacts
20: Good Characteristics due to Autism
21: Satisfaction: Given advice
22: Satisfaction: Treatment
23: Satisfaction: Medication
24: Satisfaction: Care
25: Satisfaction: Education
26: Satisfaction: Work
27: Satisfaction: Social Contacts
28: Age

Ising-Gaussian: Rewrite conditional Gaussian (1)

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

If we let $\sigma = 1$ and factor out Y_s, we get:

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

Ising-Gaussian: Rewrite conditional Gaussian (2)

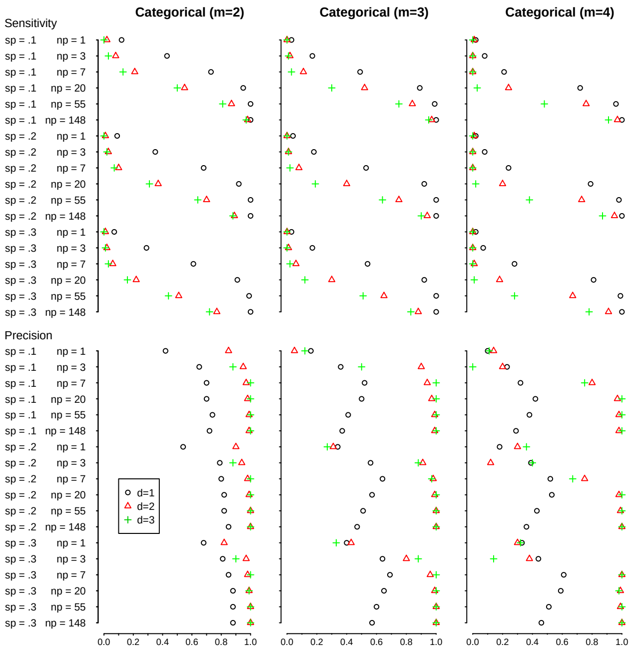
Now, if we let $\mu_s = \theta_s^y + \sum_{t \in N(s)_Y} \theta_{st}^{yy} Y_t + \sum_{r \in N(s)_Z} \theta_{sr}^{yz} Z_r$, we have

$$P(X_s|X_{\setminus s}) = \exp\left\{X_s \mu_s + \frac{X_s^2}{2} - \Phi(X_{\setminus s})\right\},$$

where $\Phi(X_{\setminus s}) = \log(\sqrt{2\pi} e^{-\frac{\mu_s^2}{2}})$. Taking $\frac{\mu_s^2}{2}$ out of the log normalization constant, with basic algebra we arrive at the well-known form of the univariate Gaussian distribution with unit variance:

$$P(X_s|X_{\setminus s}) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(X_s - \mu_s)^2}{2}\right\}$$

All results: Categorical



All results: Mixed