European Meeting of Statisticians

Palermo 22-26 July 2019

An Extension of the Censored Gaussian Graphical Lasso Estimator

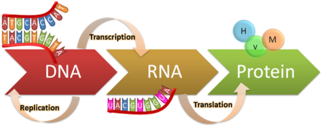
Luigi Augugliaro[1], Gianluca Sottile[1] and Veronica Vinciotti[2]

[1]Departments of Economics, Business and Statistics
University of Palermo - Italy
gianluca.sottile@unipa.it

[2]College of Engineering, Design and Physical Sciences
Brunel University - London

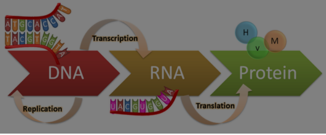
Contents

Expression Data: Complex Data from Different Platforms



- A goal in genomic is to understand interaction among genes
- These relationships are represented by a genetic network, where *nodes* represent genes and *edges* describe the interaction among them
- Typically many variables, few units ("$p \gg n$")
- A number of platforms to measure expression (mRNA) levels:
  1. quantitative real-time reverse transcription-PCR (RT-qPCR)
  2. next-generation sequencing (RNA-seq)
  3. microarray hybridization

Expression Data: Complex Data from Different Platforms



Aim: Recover/infer the underlying regulatory network from data

- A goal in genomic is to understand interaction among genes
- These relationships are represented by a genetic network, where *nodes* represent genes and *edges* describe the interaction among them
- Typically many variables, few units ("$p \gg n$")
- A number of platforms to measure expression (mRNA) levels:
  1. quantitative real-time reverse transcription-PCR (RT-qPCR)
  2. next-generation sequencing (RNA-seq)
  3. microarray hybridization

Sparse Gaussian Graphical Models

GGM in genomics:

Let $\boldsymbol{y} = (y_1, \dots, y_p)^\top$ be a p-dimensional vector of random variables

1. A normality assumption: $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with density

$$\phi(\boldsymbol{y}; \boldsymbol{\mu}, \Theta) = (2\pi)^{-p/2} |\Theta|^{1/2} \exp\{-1/2(\boldsymbol{y} - \boldsymbol{\mu})^\top \Theta (\boldsymbol{y} - \boldsymbol{\mu})\}.$$

2. A graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ is the set of p nodes (genes) and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ the set of undirected edges (genomic interactions)

Sparse Gaussian Graphical Models

GGM in genomics:

Let $\boldsymbol{y} = (y_1, \ldots, y_p)^\top$ be a p-dimensional vector of random variables

1. A normality assumption: $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with density

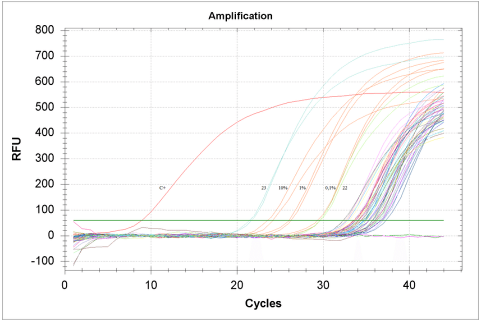
$$\phi(\boldsymbol{y}; \boldsymbol{\mu}, \Theta) = (2\pi)^{-p/2} |\Theta|^{1/2} \exp\{-1/2(\boldsymbol{y} - \boldsymbol{\mu})^\top \Theta (\boldsymbol{y} - \boldsymbol{\mu})\}.$$

2. A graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ is the set of p nodes (genes) and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ the set of undirected edges (genomic interactions)

- The precision matrix $\Theta = \boldsymbol{\Sigma}^{-1}$ provides the structure of the conditional independence graph (non-zeros $\leftrightarrow$ edges)
- Graphical lasso (glasso) estimator (Yuan and Lin, 2007)

$$\widehat{\Theta} = \arg\max_{\Theta \succ 0} \frac{1}{n} \sum_{i=1}^{n} \phi(\boldsymbol{y}_i; \boldsymbol{0}, \Theta) - \rho \sum_{h \neq k} |\theta_{hk}|. \tag{1}$$

Motivation: RT-qPCR data are censored



- Repeated cycles of DNA amplification followed by expression measurements, with a max of (typically) 40 cycles
- The cycle at which expression reaches a fixed threshold is reported
- If a gene is not expressed, the threshold is not reached after the maximum number of cycles. For this reason, the resulting data is right-censored

The Censoring Mechanism

Let $\boldsymbol{l} = (l_1, \ldots, l_p)^\top$ and $\boldsymbol{u} = (u_1, \ldots, u_p)^\top$, with $l_h < u_h$ for $h = 1, \ldots, p$ the left and right censoring values, respectively.

y_h is observed if it is inside the interval $[l_h, u_h]$, censored from below if $y_h < l_h$ or censored from above if $y_h > u_h$.

The Censoring Mechanism

Let $\boldsymbol{l} = (l_1, \ldots, l_p)^\top$ and $\boldsymbol{u} = (u_1, \ldots, u_p)^\top$, with $l_h < u_h$ for $h = 1, \ldots, p$ the left and right censoring values, respectively.

y_h is observed if it is inside the interval $[l_h, u_h]$, censored from below if $y_h < l_h$ or censored from above if $y_h > u_h$.

To obtain the joint distribution of the observed data, we follow the approach proposed by Little and Rubin (2002).

Let $\boldsymbol{r} = (r_1, \ldots, r_p)^\top$ encode the censoring patterns, with hth element

$$r_h = \begin{cases} -1, & \text{if} \quad y_h < l_h \\ 0, & \text{if} \quad l_h \leq y_h \leq u_h \\ +1, & \text{if} \quad y_h > u_h. \end{cases}$$

The Density Function under Censoring

Given a pattern of censored values, the vertex set can be partitioned into $\mathcal{V} = \boldsymbol{o} \cup \boldsymbol{c}$, where

$$\boldsymbol{o} = \{h \in \mathcal{V} : r_j = 0\} \quad \text{and} \quad \boldsymbol{c} = \{h \in \mathcal{V} : r_j \neq 0\}$$

then $\boldsymbol{y}$ can be splitted into $\boldsymbol{y}_{\boldsymbol{o}} = (y_h)_{h \in \boldsymbol{o}}$ and $\boldsymbol{y}_{\boldsymbol{c}} = (y_h)_{h \in \boldsymbol{c}}$.

The Density Function under Censoring

Given a pattern of censored values, the vertex set can be partitioned into $\mathcal{V} = \boldsymbol{o} \cup \boldsymbol{c}$, where

$$\boldsymbol{o} = \{h \in \mathcal{V} : r_j = 0\} \quad \text{and} \quad \boldsymbol{c} = \{h \in \mathcal{V} : r_j \neq 0\}$$

then $\boldsymbol{y}$ can be splitted into $\boldsymbol{y_o} = (y_h)_{h \in \boldsymbol{o}}$ and $\boldsymbol{y_c} = (y_h)_{h \in \boldsymbol{c}}$.

The joint probability distribution of $\{\boldsymbol{y}_o^\top, \boldsymbol{r}\}$ is obtained by integrating $\boldsymbol{y_c}$ out of the joint distribution of $\{\boldsymbol{y}_o^\top, \boldsymbol{y}_c^\top, \boldsymbol{r}\}$. After straightforward algebra we have:

$$\varphi(\boldsymbol{y}_o, \boldsymbol{r}; \boldsymbol{\mu}, \Theta) = \left\{ \int_{D_{\boldsymbol{c}}} \phi(\boldsymbol{y_o}, \boldsymbol{y_c}; \boldsymbol{\mu}, \Theta) \mathrm{d}\boldsymbol{y_c} \right\} I(\boldsymbol{l_o} \leq \boldsymbol{y_o} \leq \boldsymbol{u_o}), \tag{2}$$

where $D_{\boldsymbol{c}} = (-\infty, \boldsymbol{l}_{\boldsymbol{c}^-}] \times [\boldsymbol{u}_{\boldsymbol{c}^+}, +\infty)$.

Contents

Sparse Inference for cGGM

Consider a sample of n independent observations, the observed log-likelihood function can be written as

$$\ell(\boldsymbol{\mu}, \Theta) = \sum_{i=1}^{n} \log \int_{D_{c_i}} \phi(\boldsymbol{y}_{io_i}, \boldsymbol{y}_{ic_i}; \boldsymbol{\mu}, \Theta) \mathrm{d}\boldsymbol{y}_{ic_i} = \sum_{i=1}^{n} \log \varphi(\boldsymbol{y}_{i\boldsymbol{o}_i}, \boldsymbol{r}_i; \boldsymbol{\mu}, \Theta).$$

Sparse Inference for cGGM

Consider a sample of n independent observations, the observed log-likelihood function can be written as

$$\ell(\boldsymbol{\mu}, \Theta) = \sum_{i=1}^{n} \log \int_{D_{c_i}} \phi(\boldsymbol{y}_{io_i}, \boldsymbol{y}_{ic_i}; \boldsymbol{\mu}, \Theta) \mathrm{d}\boldsymbol{y}_{ic_i} = \sum_{i=1}^{n} \log \varphi(\boldsymbol{y}_{i\boldsymbol{o}_i}, \boldsymbol{r}_i; \boldsymbol{\mu}, \Theta).$$

Augugliaro et al. (2018) proposed the following estimator

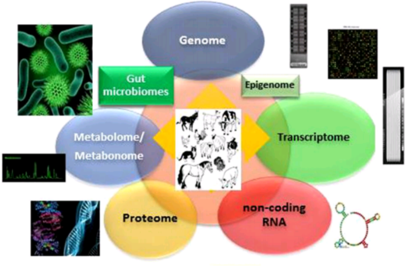
$$\{\hat{\boldsymbol{\mu}}, \widehat{\Theta}\} = \arg \max_{\boldsymbol{\mu}, \Theta \succ 0} \frac{1}{n} \ell(\boldsymbol{\mu}, \Theta) - \rho \sum_{h \neq k} |\theta_{hk}|, \qquad (3)$$

called censored Gaussian Graphical Lasso (cglasso).

Contents

Extension: Conditional Censored Gaussian Graphical Model

Genetical genomics experiments measure both genetic variants and gene expression data on the same subjects



Conditional censored GGM

$$E(\boldsymbol{y} \mid \boldsymbol{x}) = \boldsymbol{B}^\top \boldsymbol{x} \qquad (4)$$

where

- $\boldsymbol{x}_i$ is a vector of q covariates;
- $\boldsymbol{B} = \{\boldsymbol{\beta}_k\}$ is the matrix of regression coefficients;
- $\boldsymbol{\beta}_k$ is the kth column of $\boldsymbol{B}$;
- assumption: $\boldsymbol{x}$ is fully observed.

Sparse inference

Assuming that $\boldsymbol{x}$ is fully observed, the log-likelihood function of the observed data can be written as follows

$$\ell(\boldsymbol{B}, \Theta) = \sum_{i=1}^{n} \log \int_{D_{c_i}} \phi(\boldsymbol{y}_{io_i}, \boldsymbol{y}_{ic_i} \mid \boldsymbol{x}_i; \boldsymbol{B}, \Theta) \mathrm{d}\boldsymbol{y}_{ic_i}, \tag{5}$$

where

$$\phi(\boldsymbol{y} \mid \boldsymbol{x}_i; \boldsymbol{B}, \Theta) = (2\pi)^{-p/2} |\Theta|^{1/2} \exp\left\{ -\frac{1}{2} (\boldsymbol{y} - \boldsymbol{B}^\top \boldsymbol{x})^\top \Theta (\boldsymbol{y} - \boldsymbol{B}^\top \boldsymbol{x}) \right\}.$$

Sparse inference

Assuming that $\boldsymbol{x}$ is fully observed, the log-likelihood function of the observed data can be written as follows

$$\ell(\boldsymbol{B}, \Theta) = \sum_{i=1}^{n} \log \int_{D_{c_i}} \phi(\boldsymbol{y}_{io_i}, \boldsymbol{y}_{ic_i} \mid \boldsymbol{x}_i; \boldsymbol{B}, \Theta) \mathrm{d}\boldsymbol{y}_{ic_i}, \tag{5}$$

where

$$\phi(\boldsymbol{y} \mid \boldsymbol{x}_i; \boldsymbol{B}, \Theta) = (2\pi)^{-p/2} |\Theta|^{1/2} \exp\left\{ -\frac{1}{2} (\boldsymbol{y} - \boldsymbol{B}^\top \boldsymbol{x})^\top \Theta (\boldsymbol{y} - \boldsymbol{B}^\top \boldsymbol{x}) \right\}.$$

The conditional cglasso estimator is defined as:

$$\{\widehat{\boldsymbol{B}}, \widehat{\Theta}\} = \arg \max_{\boldsymbol{B}, \Theta \succ 0} \frac{1}{n} \ell(\boldsymbol{B}, \Theta) - \lambda \sum_{k=1}^{p} \theta_{kk} \|\boldsymbol{\beta}_k\|_1 - \rho \sum_{h \neq k} |\theta_{hk}|. \tag{6}$$

Contents

Sparse inference: computational aspects

A double penalized EM algorithm:

1: **repeat**

2: **E-Step**: compute the statistics $\widehat{\boldsymbol{Y}} = (\hat{y}_{i,h})$ and $\widehat{\boldsymbol{C}} = (\hat{y}_{i,hk})$

Sparse inference: computational aspects

A double penalized EM algorithm:

1: **repeat**

2:

$$\hat{y}_{i,h} = \begin{cases} y_{ih} & \text{if } r_{ih} = 0 \\ E(y_{ih} \mid \boldsymbol{y}_{ic_i} \in D_{c_i}, \boldsymbol{x}_i) & \text{otherwise,} \end{cases}$$

$$\hat{y}_{i,hk} = \begin{cases} y_{ih}y_{ik} & \text{if } r_{ih} = 0 \text{ and } r_{ik} = 0 \\ y_{ih}E(y_{ik} \mid \boldsymbol{y}_{ic_i} \in D_{c_i}, \boldsymbol{x}_i) & \text{if } r_{ih} = 0 \text{ and } r_{ik} \neq 0 \\ E(y_{ih} \mid \boldsymbol{y}_{ic_i} \in D_{c_i}, \boldsymbol{x}_i)y_{ik} & \text{if } r_{ih} \neq 0 \text{ and } r_{ik} = 0 \\ E(y_{ih}y_{ik} \mid \boldsymbol{y}_{ic_i} \in D_{c_i}, \boldsymbol{x}_i) & \text{if } r_{ih} \neq 0 \text{ and } r_{ik} \neq 0 \end{cases}$$

Sparse inference: computational aspects

A double penalized EM algorithm:

1: **repeat**

2: **E-Step**: compute the statistics $\widehat{\boldsymbol{Y}} = (\hat{y}_{i,h})$ and $\widehat{\boldsymbol{C}} = (\hat{y}_{i,hk})$

and $\widehat{\boldsymbol{S}} = n^{-1}\{\widehat{\boldsymbol{C}} - \widehat{\boldsymbol{Y}}^\top \boldsymbol{X}\widehat{\boldsymbol{B}} - \widehat{\boldsymbol{B}}^\top \boldsymbol{X}^\top \widehat{\boldsymbol{Y}} + \widehat{\boldsymbol{B}}^\top \boldsymbol{X}^\top \boldsymbol{X}\widehat{\boldsymbol{B}}\}$ (7)

Sparse inference: computational aspects

A double penalized EM algorithm:

1: **repeat**

2: **E-Step**: compute the statistics $\widehat{\boldsymbol{Y}} = (\hat{y}_{i,h})$ and $\widehat{\boldsymbol{C}} = (\hat{y}_{i,hk})$

$$\text{and } \widehat{\boldsymbol{S}} = n^{-1}\{\widehat{\boldsymbol{C}} - \widehat{\boldsymbol{Y}}^\top \boldsymbol{X}\widehat{\boldsymbol{B}} - \widehat{\boldsymbol{B}}^\top \boldsymbol{X}^\top \widehat{\boldsymbol{Y}} + \widehat{\boldsymbol{B}}^\top \boldsymbol{X}^\top \boldsymbol{X}\widehat{\boldsymbol{B}}\} \qquad (7)$$

3: **M-Step**: solve the maximization problem

$$\max_{\boldsymbol{B},\Theta \succ 0} \log\det\Theta - \mathrm{tr}\{\Theta\widehat{\boldsymbol{S}}\} - \lambda \sum_{k=1}^{p} \theta_{kk} \|\boldsymbol{\beta}_k\|_1 - \rho \sum_{h \neq k} |\theta_{hk}| \qquad (8)$$

4: **until** a convergence criterion is met

Since the penalized Q-function

$$\log\det\Theta - \operatorname{tr}\{\Theta\widehat{\boldsymbol{S}}\} - \lambda\sum_{k=1}^{p}\theta_{kk}\|\boldsymbol{\beta}_k\|_1 - \rho\sum_{h\neq k}|\theta_{hk}| \qquad (9)$$

is a bi-convex function of $\boldsymbol{B}$ and Θ (Rothman et al., 2010), its maximization can be done using the following procedure:

Since the penalized Q-function

$$\log\det\Theta - \mathrm{tr}\{\Theta\widehat{\boldsymbol{S}}\} - \lambda\sum_{k=1}^{p}\theta_{kk}\|\boldsymbol{\beta}_k\|_1 - \rho\sum_{h\neq k}|\theta_{hk}| \tag{9}$$

is a bi-convex function of $\boldsymbol{B}$ and Θ (Rothman et al., 2010), its maximization can be done using the following procedure:

1: **repeat**

2: given $\widehat{\Theta}$ and $\widehat{\boldsymbol{S}}$, solve the following problem:

$$\min_{\boldsymbol{B}}\mathrm{tr}\{\widehat{\Theta}\widehat{\boldsymbol{S}}\} + \lambda\sum_{k=1}^{p}\hat{\theta}_{kk}\|\boldsymbol{\beta}_k\|_1 \tag{10}$$

Since the penalized Q-function

$$\log\det\Theta - \mathrm{tr}\{\Theta\widehat{\boldsymbol{S}}\} - \lambda\sum_{k=1}^{p}\theta_{kk}\|\boldsymbol{\beta}_k\|_1 - \rho\sum_{h\neq k}|\theta_{hk}| \tag{9}$$

is a bi-convex function of $\boldsymbol{B}$ and Θ (Rothman et al., 2010), its maximization can be done using the following procedure:

1: **repeat**

2: given $\widehat{\Theta}$ and $\widehat{\boldsymbol{S}}$, solve the following problem:

$$\min_{\boldsymbol{B}} \mathrm{tr}\{\widehat{\Theta}\widehat{\boldsymbol{S}}\} + \lambda\sum_{k=1}^{p}\hat{\theta}_{kk}\|\boldsymbol{\beta}_k\|_1 \tag{10}$$

3: given $\widehat{\boldsymbol{B}}$, update $\widehat{\boldsymbol{S}}$ and solve the following problem:

$$\max_{\Theta\succ 0}\log\det\Theta - \mathrm{tr}\{\Theta\widehat{\boldsymbol{S}}\} - \rho\sum_{h\neq k}|\theta_{hk}| \tag{11}$$

4: **until** a convergence criterion is met

Since the penalized Q-function

$$\log \det \Theta - \mathrm{tr}\{\Theta \widehat{\boldsymbol{S}}\} - \lambda \sum_{k=1}^{p} \theta_{kk} \|\boldsymbol{\beta}_k\|_1 - \rho \sum_{h \neq k} |\theta_{hk}| \qquad (9)$$

is a bi-convex function of $\boldsymbol{B}$ and Θ (Rothman et al., 2010), its maximization

car

1:

2:

To solve Eq (10)

Since

$$\widehat{\boldsymbol{S}} = n^{-1}\{\widehat{\boldsymbol{C}} - \widehat{\boldsymbol{Y}}^\top \boldsymbol{X}\widehat{\boldsymbol{B}} - \widehat{\boldsymbol{B}}^\top \boldsymbol{X}^\top \widehat{\boldsymbol{Y}} + \widehat{\boldsymbol{B}}^\top \boldsymbol{X}^\top \boldsymbol{X}\widehat{\boldsymbol{B}}\}$$

is a function of the matrix $\boldsymbol{B}$.

Now by $\widehat{\boldsymbol{S}}(\boldsymbol{\beta}_k)$ we denote the matrix $\widehat{\boldsymbol{S}}$ seen as a function of the vector $\boldsymbol{\beta}_k$ while the other columns of the matrix $\boldsymbol{B}$ are held fixed to the current estimates.

0)

3:

1)

4: **until** a convergence criterion is met

Since the penalized Q-function

$$\log\det\Theta - \mathrm{tr}\{\Theta\widehat{\boldsymbol{S}}\} - \lambda\sum^{p}\theta_{kk}\|\boldsymbol{\beta}_k\|_1 - \rho\sum|\theta_{hk}| \tag{9}$$

is a ... on

car ...

1:

2:

... 0)

3:

... 1)

4:

Proposition

Minimization problem

$$\min_{\boldsymbol{\beta}_k}\mathrm{tr}\{\widehat{\Theta}\widehat{\boldsymbol{S}}(\boldsymbol{\beta}_k)\} + \lambda\hat{\theta}_{kk}\|\boldsymbol{\beta}_k\|_1, \tag{12}$$

is equivalent to

$$\min_{\boldsymbol{\beta}_k}\frac{1}{n}\|\widetilde{\boldsymbol{Y}_k} - \boldsymbol{X}\boldsymbol{\beta}_k\|^2 + \lambda\|\boldsymbol{\beta}_k\|_1, \tag{13}$$

where $\widetilde{\boldsymbol{Y}_k}$ is a vector with ith element

$$\tilde{y}_{i,k} = \hat{y}_{i,k} + \hat{\theta}_{kk}^{-1}\sum_{h\neq k}^{p}\hat{\theta}_{hk}\{\hat{y}_{i,h} - \boldsymbol{x}_i^\top\hat{\boldsymbol{\beta}}_h\}. \tag{14}$$

Since the penalized Q-function

p ... (9)

is a ... on

car ...

(10)

(11)

Multivariate LASSO algorithm

Problem:

$$\min_{\boldsymbol{B}} \operatorname{tr}\{\widehat{\Theta}\widehat{\boldsymbol{S}}\} + \lambda \sum_{k=1}^{p} \hat{\theta}_{kk} \|\boldsymbol{\beta}_k\|_1$$

repeat
  for $k = 1 \dots p$ **do**
    compute

$$\tilde{y}_{i,k} = \hat{y}_{i,k} + \hat{\theta}_{kk}^{-1} \sum_{h \neq k}^{p} \hat{\theta}_{hk}\{\hat{y}_{i,h} - \boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}}_h\}$$

    compute

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{k}} = \arg\min_{\boldsymbol{\beta}_k} \frac{1}{n} \|\widetilde{\boldsymbol{Y}}_k - \boldsymbol{X}\boldsymbol{\beta}_k\|^2 + \lambda \|\boldsymbol{\beta}_k\|_1$$

    update the kth column of $\widehat{\boldsymbol{B}}$ using $\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$
  end for
until convergence criterion is met

Since the penalized Q-function

$$\log\det\Theta - \mathrm{tr}\{\Theta\widehat{\boldsymbol{S}}\} - \lambda\sum_{k=1}^{p}\theta_{kk}\|\boldsymbol{\beta}_k\|_1 - \rho\sum_{h\neq k}|\theta_{hk}| \tag{9}$$

is a bi-convex function of $\boldsymbol{B}$ and Θ (Rothman et al., 2010), its maximization can be done using the following procedure:

1: **repeat**

2:

To solve Eq (11)

Standard algorithm, e.g., the block-coordinate descent algorithm proposed by Friedman.

0)

3: given $\widehat{\boldsymbol{B}}$, update $\widehat{\boldsymbol{S}}$ and solve the following problem:

$$\max_{\Theta\succ 0}\log\det\Theta - \mathrm{tr}\{\Theta\widehat{\boldsymbol{S}}\} - \rho\sum_{h\neq k}|\theta_{hk}| \tag{11}$$

4: **until** a convergence criterion is met

Contents

Simulation study: setting I

1: $\boldsymbol{X}$ is simulated from a gaussian distribution with zero expected value and sparse precision matrix (random structure)

2: for each $k = 1, \dots, p$, we randomly drawn $S_k \subset \mathcal{V}$, with $|S_k| = 2$, then:

$$\beta_{hk} \sim \mathcal{U}_{[0.3, 0.7]},$$

the remaining regression coefficients are equal to zero

3: the intercepts β_{0k} are used to set the probability of right-censoring of the first K response variables:

$$\sum_{i=1}^{n} \Pr\{y_{ik} \geq u_k\}/n = 0.4, \quad \text{for} \quad k = 1, \dots, K$$

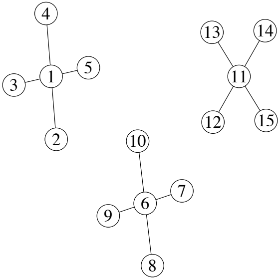
$$\sum_{i=1}^{n} \Pr\{y_{ik} \geq u_k\}/n = 10^{-6}, \quad \text{for} \quad k = K+1, \dots, p$$

Simulation study: setting II

4: the precision matrix Θ is such that $\theta_{hh} = 1$ and

$$\theta_{h(h+j)} \sim \mathcal{U}_{[0.30,0.35]}$$

with $j = 1, \ldots 4$ and $h \in \{1, 6, 11, \ldots, p\}$



Simulation study: setting III

5: finally, in each simulation run $\boldsymbol{Y}$ is simulated by the following steps:
  - the matrix $\boldsymbol{E}$ is simulated from $\mathcal{N}(\mathbf{0}; \Theta)$;
  - $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{E}$;
  - each y_{ik} greater than u_k is treated as right-censored ($u_k = 40$).

In order to evaluate the behaviour of the proposed estimator we let:

- $n = 100$;
- $K = 0.4 \times p$;
- $p, q \in \{50, 200\}$.

Competitors:

- conditional glasso estimator (Yin and Li, 2011);
- mglasso estimator (Städler and Bühlmann, 2012).

Simulation study: evaluation criteria

Sparse Recovery

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{and} \quad \text{Recall} = \frac{TP}{TP + FN}$$

The Precision-Recall curves are obtained fixing a tuning parameter, for example ρ, and varying the remaining tuning parameter.

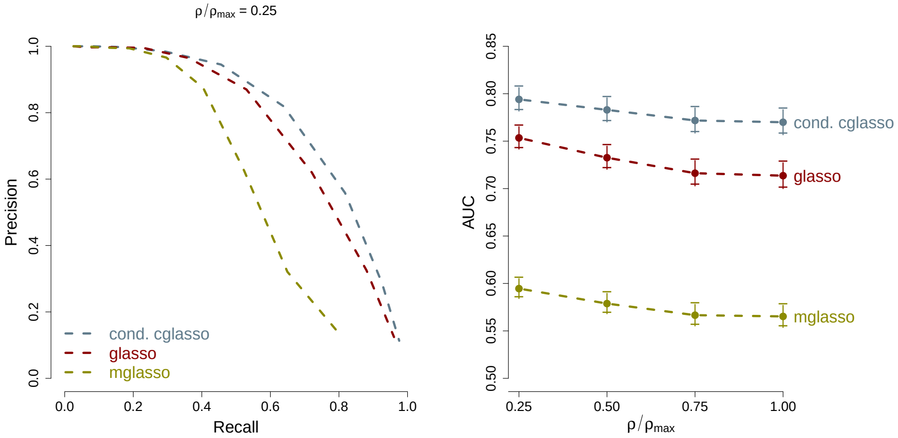
Mean Square Error

$$\begin{aligned} \text{MSE}(\widehat{\boldsymbol{B}}) &= E\{\|\widehat{\boldsymbol{B}} - \boldsymbol{B}^\star\|_F^2\} \\ \text{MSE}(\widehat{\Theta}) &= E\{\|\widehat{\Theta} - \Theta^\star\|_F^2\} \end{aligned}$$

where $\| \cdot \|_F^2$ denotes the Frobenius norm.

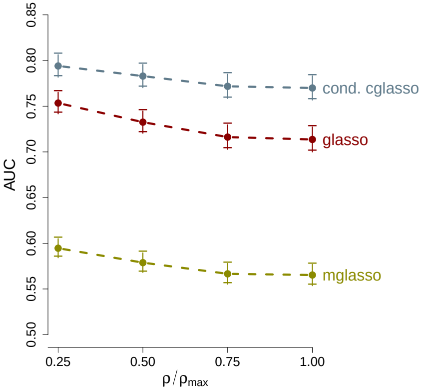
Simulation study: results on $\widehat{\boldsymbol{B}}$ (Sparsity Recovery)

Setting: $n = 100, p = 50, q = 50$

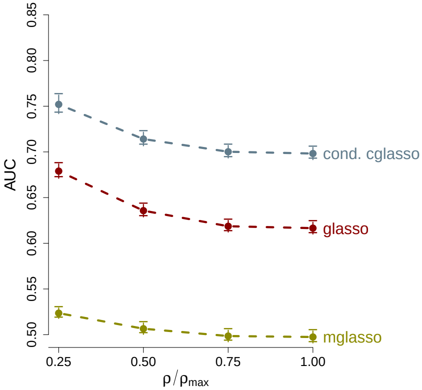


Simulation study: results on $\widehat{\boldsymbol{B}}$ (Sparsity Recovery)

Setting: $n = 100$



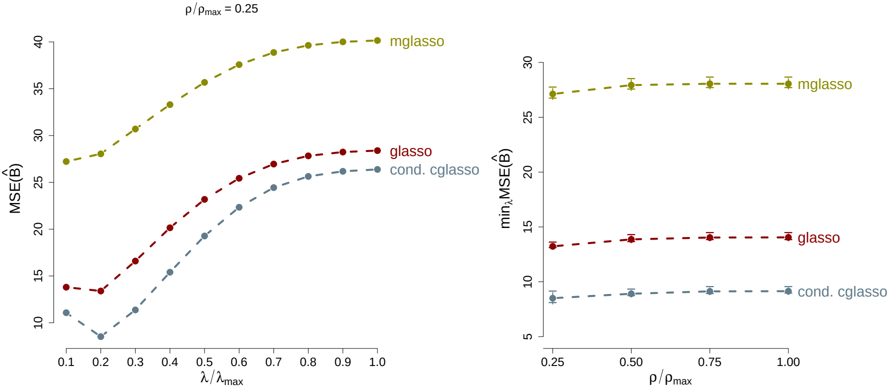
(a) $p = 50, q = 50$



(b) $p = 200, q = 200$

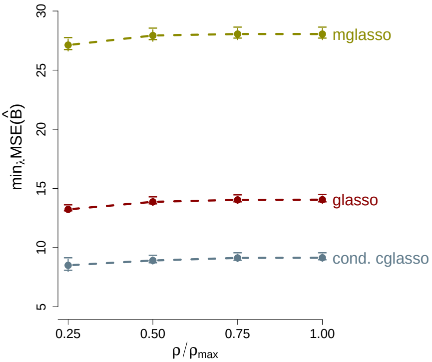
Simulation study: results on $\widehat{\boldsymbol{B}}$ (Mean Square Error)

Setting: $n = 100, p = 50, q = 50$

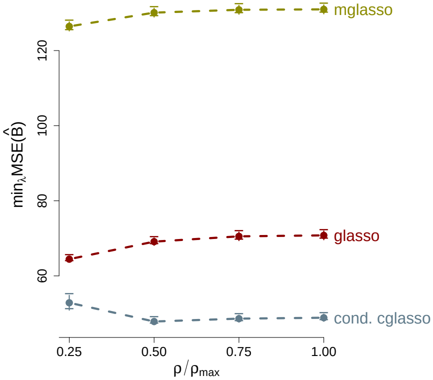


Simulation study: results on $\widehat{B}$ (Mean Square Error)

Setting: $n = 100$



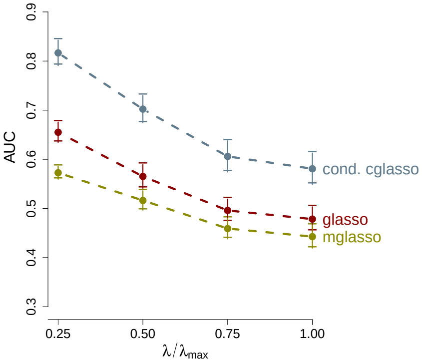
(c) $p = 50, q = 50$



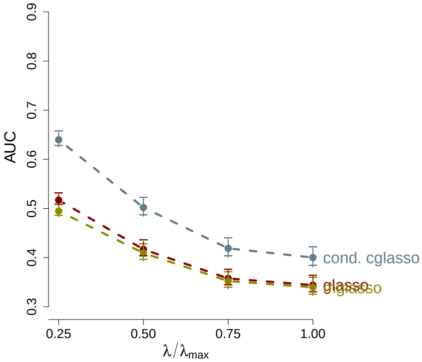
(d) $p = 200, q = 200$

Simulation study: results on $\widehat{\Theta}$ (Sparsity Recovery)

Setting: $n = 100$



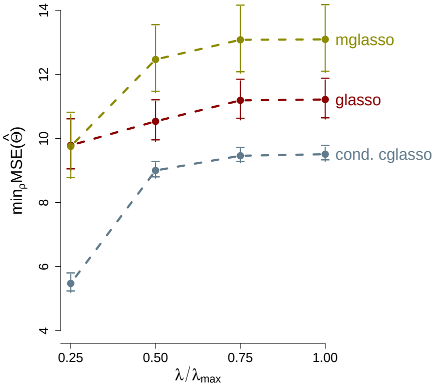
(e) $p = 50, q = 50$



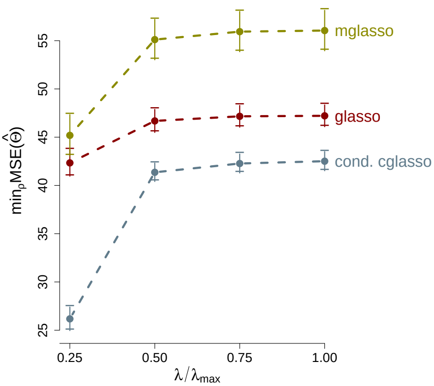
(f) $p = 200, q = 200$

Simulation study: results on $\widehat{\Theta}$ (Mean Square Error)

Setting: $n = 100$



(g) $p = 50, q = 50$



(h) $p = 200, q = 200$

Conclusions and future work

Summary:

- a new method for sparse conditional Gaussian graphical models with censored data;
- an efficient EM algorithm.

Conclusions and future work

Summary:

- a new method for sparse conditional Gaussian graphical models with censored data;
- an efficient EM algorithm.

Future work:

- censored covariates;
- a measure of goodness-of-fit;
- `cglasso` package.

References

Augugliaro, L., Mineo, A.M. and Wit, E.C. (2016). *ℓ_1-Penalized methods in high-dimensional Gaussian Markov random fields*. In: Computational Network Analysis with R: Applications in Biology, Medicine, and Chemistry. Wiley-VCH, pp. 201-267.

Augugliaro, L., Abbruzzo, A. and Vinciotti, V. (2018). *ℓ_1-Penalized censored Gaussian graphical model*. Biostatistics. kxy043 `https://doi.org/10.1093/biostatistics/kxy043`

Friedman, J.H., Hastie T. and Tibshirani, R. (2008). *Sparce inverse covariance estimation with the graphical lasso*. Biostatistics, 9(3):432–441.

Rothman, A. J., Levina, E. and Zhu J. (2010). *Sparse multivariate regression with covariance estimation. Journal of Computational and Graphical Statistics*, 19(4):947–962.

Städler, N. and Bühlmann, P. (2012). *Missing values: sparse inverse covariance estimation and an extension to sparse regression*. Statistics and Computing, 22(1):219–235.

Yin, J. and Li, H. (2011). *A sparse conditional Gaussian graphical model for analysis of genetical genomics data*. The Annals of Applied Statistics, 5(4):2630–2650.


aitäh
spas
nuhun
stuutiyi
nandri
supas
akiba
tack
saðol
tänan
merci
buznyg
paldies
matondi
taiku
wado
matondo
marahaba
tashakor
thank you
grazie
mahalo
mèsi
tenki
saha
waybale
rahmat
tanmirt
yekeniele
obrigado
meharbani
ahsante
dankewol
talofa
obrigada
manjuthe
bayarlalaa
hvala
chokrane
omol
gracias
dankegon
kiitos
shukran
ngiyabonga
takk
soolong
dakujem
aabar
shukriyaa
salamat
manana
modupe
dziekuje
zikomo
miigwech
efharisto
blagodaram
cheers
köszönöm
gràcies
blagodaria
danki
multumesc
misaotra
waita
gratzias
murakoze
shakkran
dziakuju
spasiba
barkal
sobodi
menlau
grassie
shukria
welalin
danke
dėkoju
skee
arigato
merkzi
vinaka
kinisou
sulpáy
trugéré
tanemirt
dekuji
maururu
bedankt
madlobt