Latent Variable Models for the Analysis, Visualization and Prediction of Network and Nodal Attribute Data

Isabella Gollini

School of Engineering
University of Bristol
isabella.gollini@bristol.ac.uk

January 24th, 2014

Joint work with Prof. Brendan Murphy (University College Dublin)

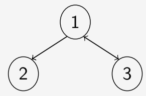
About Me

- With Jonty: Probabilistic methods for uncertainty assessment and quantification in natural hazards (floods, volcanoes, and earthquakes etc.).
- Models to cluster binary data with complex dependence structure
  - Gollini, I., and Murphy, T.B., (2013) "Mixture of Latent Trait Analyzers for Model-Based Clustering of Categorical Data", *Statistics and Computing*.
- Models for network data
- Use of Variational methods for fast approximate inference.

Outline

Notation

- N number of nodes of an observed network



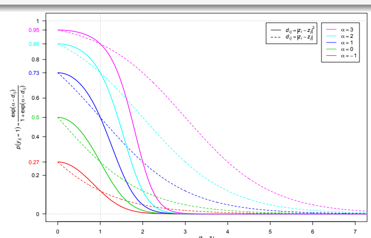
- $\mathbf{Y}$ $(N \times N)$ adjacency matrix
  $$y_{ij} = \begin{cases} 1 & \text{if there is an edge between node } i \text{ and } j \\ 0 & \text{otherwise} \end{cases}$$
- $\mathbf{X}$ $(N \times M)$ matrix of M nodal attributes.
- $\mathbf{z}_n \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ D dimensional continuous latent variable

Latent Space Model (LSM) for Networks

- Hoff *et al.* (2002) introduced a model that assumes that each node n has an unknown position $\mathbf{z}_n$ in a D-dim *Euclidean latent space*.
  $$p(\mathbf{Y}|\mathbf{Z},\alpha) = \prod_{i \neq j}^{N} p(y_{ij}|\mathbf{z}_i,\mathbf{z}_j,\alpha) = \prod_{i \neq j}^{N} \frac{\exp(\alpha - |\mathbf{z}_i - \mathbf{z}_j|^2)^{y_{ij}}}{1+\exp(\alpha - |\mathbf{z}_i - \mathbf{z}_j|^2)}$$
  with $p(\alpha) = \mathcal{N}(\xi, \psi^2)$, $p(\mathbf{z}_n) \overset{iid}{=} \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_D)$ and σ^2, ξ, ψ^2 are fixed parameters.
- The posterior distribution *cannot* be calculated analytically.
- NOTE: We propose to use is the *Squared Euclidean Distance.*

Why Squared Euclidean Distance?

- It requires less approximation to be made in the estimation procedure.
- It allows to visualize more clearly the presence of potential clusters, giving a higher probability of a link between two close nodes in the latent space and lower probability to two nodes lying far away from each other.



LSM for Networks – Variational Approach

- We fit the model using a *Variational inference approach* that is considerably quicker but less accurate than MCMC.

- The posterior probability of the unknown $(\mathbf{Z}, \alpha)$ is:

$$p(\mathbf{Z}, \alpha|\mathbf{Y}) = p(\mathbf{Y}|\mathbf{Z}, \alpha)p(\alpha)\prod_{n=1}^{N} p(\mathbf{z}_n) \times C$$

where C is the unknown normalising constant

- We propose a variational posterior $q(\mathbf{Z}, \alpha|\mathbf{Y})$ introducing variational parameters $\tilde{\xi}, \tilde{\psi}^2, \tilde{\mathbf{z}}_n, \tilde{\mathbf{\Sigma}}$:

$$q(\mathbf{Z}, \alpha|\mathbf{Y}) = q(\alpha)\prod_{n=1}^{N} q(\mathbf{z}_n)$$

where $q(\alpha) = \mathcal{N}(\tilde{\xi}, \tilde{\psi}^2)$ and $q(\mathbf{z}_n) = \mathcal{N}(\tilde{\mathbf{z}}_n, \tilde{\mathbf{\Sigma}})$.

Variational Approach

- The basic idea behind the variational approach is to find a lower bound of the log marginal likelihood $\log p(\mathbf{Y})$ by introducing the variational posterior distribution $q(\mathbf{Z}, \alpha|\mathbf{Y})$.

- This approach leads to minimize the Kulback-Leibler divergence between the variational posterior $q(\mathbf{Z}, \alpha|\mathbf{Y})$ and the true posterior $p(\mathbf{Z}, \alpha|\mathbf{Y})$:

$$\begin{aligned}
\mathrm{KL}[q(\mathbf{Z}, \alpha|\mathbf{Y})||p(\mathbf{Z}, \alpha|\mathbf{Y})] &= -\int q(\mathbf{Z}, \alpha|\mathbf{Y}) \log \frac{p(\mathbf{Z}, \alpha|\mathbf{Y})}{q(\mathbf{Z}, \alpha|\mathbf{Y})} d(\mathbf{Z}, \alpha) \\
&= \int q(\mathbf{Z}, \alpha|\mathbf{Y}) \log \frac{p(\mathbf{Y}, \mathbf{Z}, \alpha)}{p(\mathbf{Y})q(\mathbf{Z}, \alpha|\mathbf{Y})} d(\mathbf{Z}, \alpha) \\
&= \int q(\mathbf{Z}, \alpha|\mathbf{Y}) \log \frac{p(\mathbf{Y}, \mathbf{Z}, \alpha)}{q(\mathbf{Z}, \alpha|\mathbf{Y})} d(\mathbf{Z}, \alpha) - \log p(\mathbf{Y})
\end{aligned}$$

Variational Approach

- $\mathrm{KL}[q(\mathbf{Z}, \alpha|\mathbf{Y})||p(\mathbf{Z}, \alpha|\mathbf{Y})]$ divergence can be written as:

$$\begin{aligned}
\mathrm{KL}[q(\mathbf{Z}, \alpha|\mathbf{Y})||p(\mathbf{Z}, \alpha|\mathbf{Y})] = \mathrm{KL}[q(\alpha)||p(\alpha)] + \sum_{i=1}^{N} \mathrm{KL}[q(\mathbf{z}_i)||p(\mathbf{z}_i)] \\
- \mathbb{E}_{q(\mathbf{Z}, \alpha|\mathbf{Y})}[\log(p(\mathbf{Y}|\mathbf{Z}, \alpha))]
\end{aligned}$$

$\mathbb{E}_{q(\mathbf{Z}, \alpha|\mathbf{Y})}[\log(p(\mathbf{Y}|\mathbf{Z}, \alpha))]$ is approximated using the Jensen's inequality:

$$\begin{aligned}
\mathbb{E}_{q(\mathbf{Z}, \alpha|\mathbf{Y})}[\log(p(\mathbf{Y}|\mathbf{Z}, \alpha))] &= \sum_{i \neq j}^{N} y_{ij}\mathbb{E}_{q(\mathbf{Z}, \alpha|\mathbf{Y})}[\alpha - |\mathbf{z}_i - \mathbf{z}_j|^2] \\
&\quad - \mathbb{E}_{q(\mathbf{Z}, \alpha|\mathbf{Y})}[\log(1 + \exp(\alpha - |\mathbf{z}_i - \mathbf{z}_j|^2))] \\
&\leq \sum_{i \neq j}^{N} y_{ij}(\mathbb{E}_{q(\mathbf{Z}, \alpha|\mathbf{Y})}[\alpha - |\mathbf{z}_i - \mathbf{z}_j|^2]) \\
&\quad - \log(1 + \mathbb{E}_{q(\mathbf{Z}, \alpha|\mathbf{Y})}[\exp(\alpha - |\mathbf{z}_i - \mathbf{z}_j|^2)])
\end{aligned}$$

LSM for Networks – Variational Approach – EM Algorithm

- The EM algorithm at each $(i+1)th$ iteration:

E-Step Estimate $\tilde{\mathbf{z}}_n^{(i+1)}$ and $\tilde{\mathbf{\Sigma}}^{(i+1)}$:

$$\mathcal{Q}(\Theta_{LSM}; \Theta_{LSM}^{(i)}) = \mathrm{KL}[q(\mathbf{Z}, \alpha|\mathbf{Y})||p(\mathbf{Z}, \alpha|\mathbf{Y})]$$

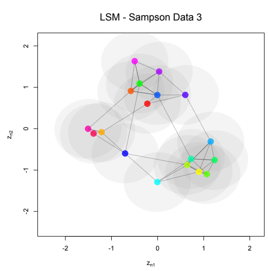
where $\Theta_{LSM} = (\tilde{\xi}, \tilde{\psi}^2)$.

M-Step Estimate $\tilde{\xi}$ and $\tilde{\psi}^2$:

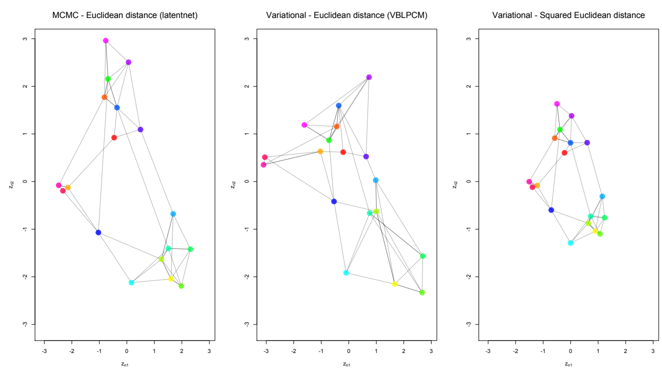
$$\Theta_{LSM}^{(i+1)} = \operatorname{argmax} \mathcal{Q}(\Theta_{LSM}; \Theta_{LSM}^{(i)})$$

LSM – Monks Network

- Sampson (1969) recorded the social interactions among a group of $N = 18$ monks while being a resident in a New England monastery.
- The directed links of the network represent the liking relationships.



Comparison of Estimation Methods and Distance Metrics



Variational Inference VS MCMC

- Closed form posteriors
- Far faster than MCMC based methods
- In the absence of posterior dependence, the lower bound would match the log likelihood.
- As long as the posterior dependence is weak, the VA may be useful:
  - Larger networks
  - For starting point of MCMC algorithms
  - To explore the model space.

- Underestimates variances
- Difficult to assess how tight the lower bound is.
- Sensitive to starting values (local minima)

Network and Nodal Attributes

- The classical approach to incorporate nodal attributes in the LSM is:

$$p(\mathbf{Y}|\mathbf{Z},\alpha,\mathbf{X},\beta) = \prod_{i\neq j}^{N} \frac{\exp(\alpha+\beta^T\mathbf{x}_{ij}-|\mathbf{z}_i-\mathbf{z}_j|^2)^{y_{ij}}}{1+\exp(\alpha+\beta^T\mathbf{x}_{ij}-|\mathbf{z}_i-\mathbf{z}_j|^2)}$$

- β and $\mathbf{x}_{ij}$ are vectors of length M.

- This LSM contains only link covariate information $\mathbf{x}_{ij}$ so it is not designed to deal with nodal attributes directly.

- This model assumes that the probability of a link depends on the nodal attributes (*social selection*)
- Sometimes the nodal attributes depend on the network links (*social influence*).
- We present a model where the network and the nodal attributes data mutually depend on each other.

Factor Analysis (FA) for Nodal Attributes

- Factor analysis (FA) (Spearman, 1904) is a useful technique to visualize continuous data, reducing the data dimensionality from M to D (where $D \ll M$) in order to explain the variability expressed by the correlation within the data.

- FA assumes that there is a continuous latent variable $\mathbf{z}_n$ underlying the behavior of the continuous response variables given by an observation $\mathbf{x}_n$.

Factor Analysis (FA) for Nodal Attributes

$$\mathbf{z}_n \overset{iid}{\sim} \mathcal{N}(\mathbf{0},\sigma^2\mathbf{I}), \qquad \text{and} \qquad \varepsilon_n \overset{iid}{\sim} \mathcal{N}(\mathbf{0},\mathbf{\Psi}),$$

where $\mathbf{\Psi} = \mathrm{diag}(\psi_1^2,\ldots,\psi_M^2)$, and

$$\mathbf{x}_n = \mu + \mathbf{\Lambda}\mathbf{z}_n + \varepsilon_n$$

So,

$$p(\mathbf{x}_n|\mathbf{z}_n) \sim \mathcal{N}(\mu, (\mathbf{\Lambda}\sigma)(\mathbf{\Lambda}\sigma)^T + \mathbf{\Psi})$$

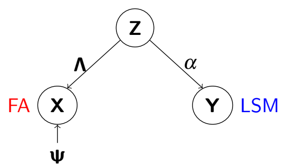
- The EM algorithm is used to find maximum likelihood estimate.

$$p(\mathbf{z}_n|\mathbf{x}_n) \sim \mathcal{N}(\hat{\mathbf{z}}_n, \hat{\mathbf{\Sigma}})$$

- Everything can be calculated analytically in closed form.

The joint model for Network and Nodal Attributes

- The probability of a node being connected with other nodes and the behaviour of nodal attributes are explained by the same latent variable.
- A continuous latent variable $\mathbf{z}_n \sim \mathcal{N}(0,\sigma^2\mathbf{I}_D)$ summarizes the information given by both the network and the nodal attributes.
- Network $\mathbf{Y}$ and nodal attributes $\mathbf{x}_n$ are independent given the latent variable $\mathbf{z}_n$.



The joint model for Network and Nodal Attributes – Fit the model

- We assume that: $p(\mathbf{z}_n) \sim \mathcal{N}(\mathbf{0},\sigma^2\mathbf{I})$.
- The network data are modeled via LSM: $p(\mathbf{z}_n|\mathbf{Y}) \sim \mathcal{N}(\tilde{\mathbf{z}}_n, \tilde{\mathbf{\Sigma}})$.
- The nodal attributes are modeled via FA: $p(\mathbf{z}_n|\mathbf{x}_n) \sim \mathcal{N}(\hat{\mathbf{z}}_n, \hat{\mathbf{\Sigma}})$.

- Joint model:

$$p(\mathbf{z}_n|\mathbf{Y},\mathbf{x}_n) \propto \frac{p(\mathbf{z}_n|\mathbf{Y})p(\mathbf{z}_n|\mathbf{x}_n)}{p(\mathbf{z}_n)} \propto \mathcal{N}(\bar{\mathbf{z}}_n, \bar{\mathbf{\Sigma}})$$

where

$$\bar{\mathbf{\Sigma}} = \left[\tilde{\mathbf{\Sigma}}^{-1} + \hat{\mathbf{\Sigma}}^{-1} - \frac{1}{\sigma^2}\mathbf{I}_D\right]^{-1} \quad \text{and} \quad \bar{\mathbf{z}}_n = \bar{\mathbf{\Sigma}}\left[\tilde{\mathbf{\Sigma}}^{-1}\tilde{\mathbf{z}}_n + \hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{z}}_n\right]$$

The joint model for Network and Nodal Attributes– EM Algorithm

E-Step Estimate $\bar{\boldsymbol{\Sigma}}^{(i+1)}$ and $\bar{\mathbf{z}}_n^{(i+1)}$

$$\mathcal{Q}(\Theta_{LSM},\Theta_{FA};\Theta^{(i)}_{LSM},\Theta^{(i)}_{FA}) = $$
$$= \mathbb{E}_{p(\mathbf{Z}|\mathbf{Y},\mathbf{X};\Theta^{(i)}_{LSM},\Theta^{(i)}_{FA})}[\log(p(\mathbf{Y},\mathbf{Z}|\Theta_{LSM}))] +$$
$$+ \mathbb{E}_{p(\mathbf{Z}|\mathbf{Y},\mathbf{X};\Theta^{(i)}_{LSM},\Theta^{(i)}_{FA})}[\log(p(\mathbf{X},\mathbf{Z}|\Theta_{FA}))]$$

therefore,

$$\bar{\boldsymbol{\Sigma}}^{(i+1)} = \left[[\tilde{\boldsymbol{\Sigma}}^{(i+1)}]^{-1} + [\hat{\boldsymbol{\Sigma}}^{(i+1)}]^{-1} - \frac{1}{\sigma^2}\mathbf{I}_D\right]^{-1}$$

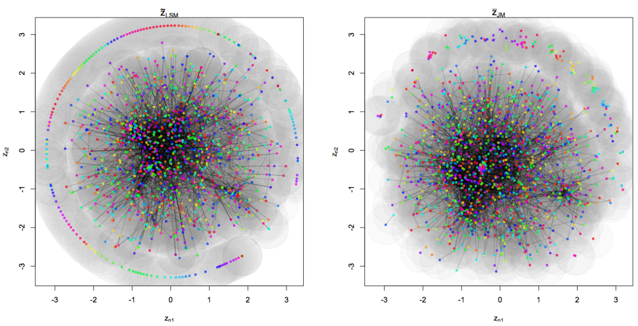
$$\bar{\mathbf{z}}_n^{(i+1)} = \bar{\boldsymbol{\Sigma}}^{(i+1)}\left[[\tilde{\boldsymbol{\Sigma}}^{(i+1)}]^{-1}\tilde{\mathbf{z}}_n^{(i+1)} + [\hat{\boldsymbol{\Sigma}}^{(i+1)}]^{-1}\hat{\mathbf{z}}_n^{(i+1)}\right]$$

M-Step Update
$$(\Theta^{(i+1)}_{LSM},\Theta^{(i+1)}_{FA}) = \text{argmax}\ \mathcal{Q}(\Theta_{LSM},\Theta_{FA};\Theta^{(i)}_{LSM},\Theta^{(i)}_{FA})$$

The joint model for Network and Nodal Attributes – Yeast Proteins

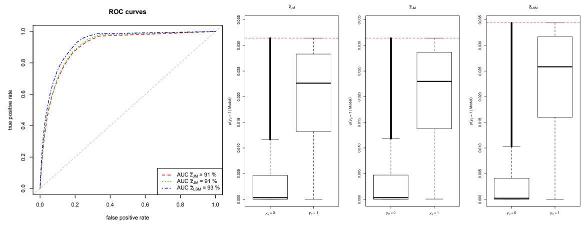
- The network is composed by N=1500 nodes representing the interaction between Saccharomyces cerevisiae (yeast) proteins.
- The nodal attributes consist of expression levels during yeast sporulation from M=80 experiments with Saccharomyces cerevisiae proteins.
- Factor analysis is an appropriate tool to visualize the M-dimensional expression data in a low dimensional latent space.
- The fixed parameters:
  - $\mathbf{z}_n \sim \mathcal{N}(\mathbf{0},\mathbf{I}_2)$
  - $\xi = 0$
  - $\psi^2 = 2$

The joint model for Network and Nodal Attributes – Results



LSM positions (left) and LSJM positions (right).

The joint model for Network and Nodal Attributes – Performance



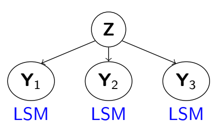
ROC curve (left) and Boxplot (right) of the estimated probabilities of a link for the true negatives and true positives.

Multiple Network Views

- In many applications the behaviour of the nodes is strongly shaped by the complex relation of many interactions.
- *Longitudinal networks*: the links represent the same relation at different time points.
- *Multiplex networks*: the links come from different kind of relations (eg genetic and physical etc.)

Joint Modelling of Multiple Network Views

- We have K networks on the same N nodes. We propose a model that merges the information given by all these networks.
- A continuous latent variable $\mathbf{z}_n \sim \mathcal{N}(0,\sigma^2\mathbf{I}_D)$ identifies the position of node n in a D-dimensional latent space.



Joint Modelling of Multiple Network Views – Model

- The probability of a link depends on the distance between two nodes in the latent space.

$$p(\mathbf{Y}_1,\dots,\mathbf{Y}_K|\mathbf{Z},\alpha) = \prod_{k=1}^{K}\prod_{i\neq j}^{N} \frac{\exp(\alpha_k - |\mathbf{z}_i-\mathbf{z}_j|^2)^{y_{ijk}}}{1+\exp(\alpha_k - |\mathbf{z}_i-\mathbf{z}_j|^2)}$$

- Variational Approach $k=1,\dots,K$: $p(\mathbf{z}_n|\mathbf{Y}_k)\sim\mathcal{N}(\tilde{\mathbf{z}}_{nk},\tilde{\boldsymbol{\Sigma}}_k)$.
- Joining the two models:

$$p(\mathbf{z}_n|\mathbf{Y}_1,\dots,\mathbf{Y}_K;\Theta_1,\dots,\Theta_K) \propto \frac{\prod_{k=1}^{K} p(\mathbf{z}_n|\mathbf{Y}_k;\Theta_k)}{p(\mathbf{z}_n)^{K-1}} \propto \mathcal{N}(\bar{\mathbf{z}}_n,\bar{\boldsymbol{\Sigma}})$$

- where

$$\bar{\boldsymbol{\Sigma}} = \left[\sum_{k=1}^{K}\tilde{\boldsymbol{\Sigma}}_k^{-1} - \frac{K-1}{\sigma^2}\mathbf{I}_D\right]^{-1} \text{ and } \bar{\mathbf{z}}_n = \bar{\boldsymbol{\Sigma}}\left[\sum_{k=1}^{K}\tilde{\boldsymbol{\Sigma}}_k^{-1}\tilde{\mathbf{z}}_{nk}\right]$$

Joint Modelling of Multiple Network Views – EM Algorithm

E-Step Estimate the parameters of the joint posterior distribution $\bar{\boldsymbol{\Sigma}}^{(i+1)}$ and $\bar{\mathbf{z}}_n^{(i+1)}$:

$$\mathcal{Q}(\Theta_1,\dots,\Theta_K;\Theta_1^{(i)},\dots,\Theta_K^{(i)}) =$$

$$= \sum_{k=1}^{K}\mathbb{E}_{p(\mathbf{Z}|\mathbf{Y}_1,\dots,\mathbf{Y}_K;\Theta_1^{(i)},\dots,\Theta_K^{(i)})}[\log(p(\mathbf{Y}_k,\mathbf{Z}|\Theta_k))]$$

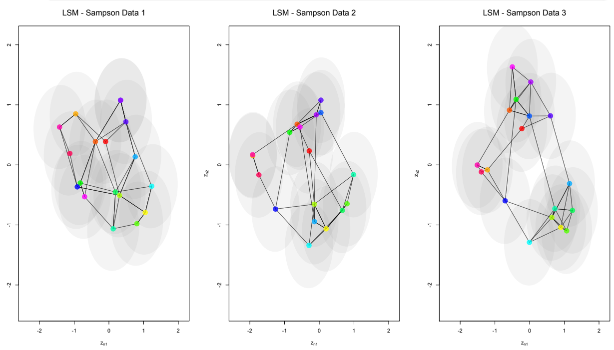
- We estimate the parameters $\tilde{\mathbf{z}}_{nk},\tilde{\boldsymbol{\Sigma}}_k$ of the posterior distribution $p(\mathbf{z}_n|\mathbf{Y}_k;\Theta_k)$ given each network k separately.
- We merge these estimates to find joint posterior distribution of the latent positions $\mathcal{N}(\bar{\mathbf{z}}_n,\bar{\boldsymbol{\Sigma}})$.

M-Step Update the variational model parameters $\tilde{\xi}_1,\dots,\tilde{\xi}_K,\tilde{\psi}_1^2,\dots,\tilde{\psi}_K^2$:

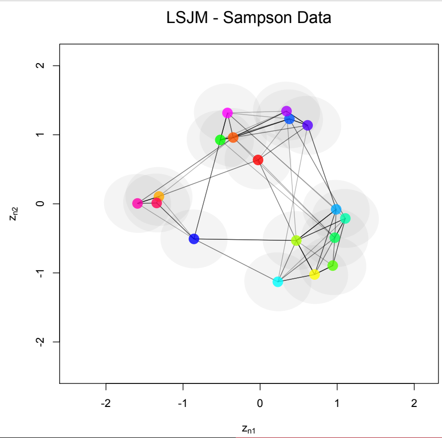
$$(\Theta_1^{(i+1)},\dots,\Theta_K^{(i+1)}) = \operatorname{argmax}\ \mathcal{Q}(\Theta_1,\dots,\Theta_K;\Theta_1^{(i)},\dots,\Theta_K^{(i)}).$$

LSJM – Monks Network (cont'd)

- We analyze the networks of liking relationship at $K=3$ time points fitting the LSM to each network separately.



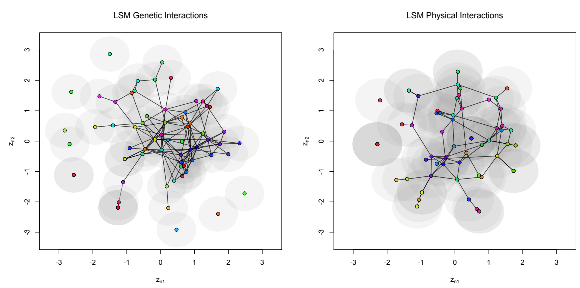
LSJM – Monks Network – LSJM positions



LSJM – Protein-Protein Interactions

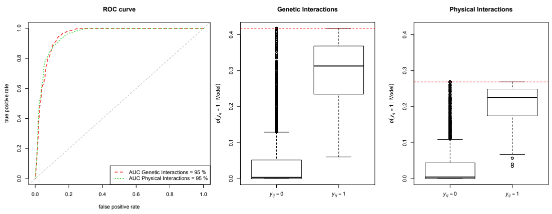
- $K=2$ undirected networks formed by genetic and physical protein-protein interactions between $N=67$ Saccharomyces cerevisiae proteins.
- The complex relational structure of this dataset has led to implementation of models aiming at describing the functional relationships between the observations.
- The data were downloaded from the Biological General Repository for Interaction Datasets (BioGRID) database.

LSJM – Protein-Protein Interactions – LSM positions



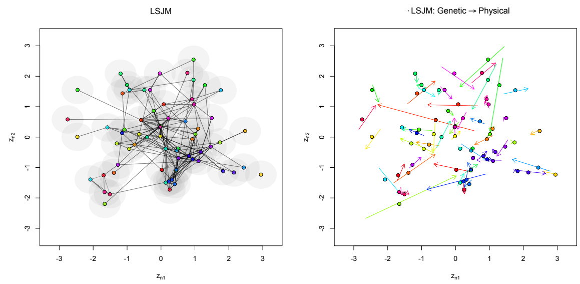
Latent posterior distributions fitting the LSM for the two networks separately.

LSJM – Protein-Protein Interactions – LSM ROC and BOX plots



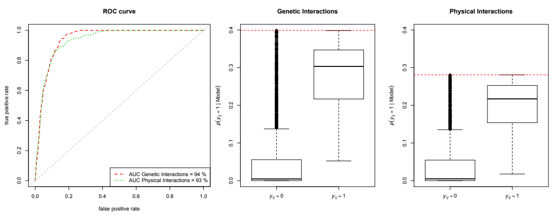
ROC curves and Boxplots of the estimated probabilities of a link for the true negatives and true positives.

LSJM – Protein-Protein Interactions – LSJM positions



- Left: $p(\mathbf{z}_n|\mathbf{Y}_1,\dots,\mathbf{Y}_K;\Theta_1,\dots,\Theta_K) \propto \mathcal{N}(\bar{\mathbf{z}}_n,\bar{\boldsymbol{\Sigma}})$ fitting the LSJM
- Right: the dots represent the overall positions $\bar{\mathbf{z}}_n$ and the arrows connect the estimated position under each model $p(\mathbf{z}_n|\mathbf{Y}_k;\Theta_k) = \mathcal{N}(\tilde{\mathbf{z}}_{nk},\tilde{\boldsymbol{\Sigma}}_k)$.

LSJM – Protein-Protein Interactions – LSJM ROC and BOX plots



ROC curves and Boxplots of the estimated probabilities of a link for the true negatives and true positives.

LSJM – Missing Links and Missing Nodes

- Missing (unobserved) links can be easily managed by the LSJM using the information given by all the network views.

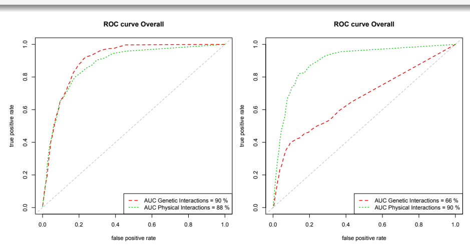
- To estimate the probability of the presence or absence of an edge we employ the posterior mean of the α_k and of the latent positions so that we get the following equation:

$$y^*_{ijk} = p(y_{ijk}=1|\bar{\mathbf{z}}_i,\bar{\mathbf{z}}_j,\tilde{\xi}_k) = \frac{\exp(\tilde{\xi}_k - |\bar{\mathbf{z}}_i-\bar{\mathbf{z}}_j|^2)}{1+\exp(\tilde{\xi}_k - |\bar{\mathbf{z}}_i-\bar{\mathbf{z}}_j|^2)}.$$

- If we want to infer whether to assign $y_{ijk}=1$ or not, we need to introduce a threshold τ_k, and let $y_{ijk}=1$ if $p(y_{ijk}=1|\tilde{\mathbf{z}}_{ik},\tilde{\mathbf{z}}_{jk},\tilde{\xi}_k) > \tau_k$.

LSJM – Protein-Protein Interactions – Missing Data

- To evaluate the link prediction we applied a 10-fold cross validation setting the 10% of the links to be missing at each time point.



ROC curves fitting a LSJM (left) and 2 single LSM (right)

LSJM – Protein-Protein Interactions – Missing Data

- Missing Links (10-fold cross validation):
  - LSJM: misclassification rate of 9% for the genetic interaction network, and 6% for the physical interaction network.
  - LSM: misclassification rate of 18% for the genetic interaction network, and 7% for the physical interaction network.

- Missing Nodes (10-fold cross validation):
  - LSJM: misclassification rate of 24% for the genetic interactions dataset and 20% for the physical interaction network.
  - LSM: *useless* since it would locate the nodes only relying on the prior information.

- Try to improve the predictions using a higher dimension for the latent variables.

Conclusions

The joint models are particularly useful to:

- Locate unconnected nodes/subgraphs in the latent space.
- Estimate missing links.
- Wide range of applications

Variational Bayes allows to deal with networks of thousands of nodes.

Possible extentions:

- Joint models for directed networks using the inner product instead of Euclidean distance.
- Joint models for categorical nodal attributes (LTA instead of FA).
- Joint models with clusters (LPCM, MFA, MLTA).
- Beyond binary networks: Rank and Count data.

References

- Hoff, P. D., Raftery, A. E., and Handcock, M. S. (2002). "Latent space approaches to social network analysis." *Journal of the american Statistical association*, 97 (460), 1090–1098.
- Salter-Townshend, M., White, A., Gollini, I., and Murphy, T.B. (2012). "Review of Statistical Network Analysis: Models, Algorithms and Software", *Statistical Analysis and Data Mining*, 5 (4), 243–264.
- Gollini, I., and Murphy, T.B. (2013). "Joint Modelling of Multiple Network Views", arXiv:1301.3759