IV. Analyse de réseaux biologiques

Catherine Matias

CNRS - Laboratoire de Probabilités et Modèles Aléatoires, Paris
catherine.matias@math.cnrs.fr
http://cmatias.perso.math.cnrs.fr/

ENSAE - 2014/2015
Some statistical networks models
   The 'most famous' ones
   Exponential random graphs
   (Overlapping) Stochastic block models
   Latent space models

Analyzing networks: (probabilistic) node clustering

Applications to biological networks
Outline

Some statistical networks models
  The ‘most famous’ ones
  Exponential random graphs
  (Overlapping) Stochastic block models
  Latent space models

Analyzing networks: (probabilistic) node clustering

Applications to biological networks
Notation

- $G = (V, E)$ is a graph with nodes set $V = \{1, \ldots, n\}$ and edges set $E$.
- For any $i, j \in V$, random variable $X_{ij} = 1\{(i, j) \in E\}$ is the edge indicator (binary graph). Sometimes, we will consider weighted graphs and then $X_{ij} \in \mathbb{R}$ is a weight on edge $(i, j) \in E$.
- $X = (X_{ij})_{1 \leq i,j \leq n}$ is the adjacency matrix of the graph.
- Graphs may be undirected ($X_{ij} = X_{ji}, 1 \leq i \leq j \leq n$) or directed ($X_{ij} \neq X_{ji}, 1 \leq i, j \leq n$). They may admit self-loops (random variables $X_{ii}$) or not (then set $X_{ii} = 0$).
- For undirected graphs, $D_i = \sum_{j \neq i} X_{ij}$ is the degree of node $i$.
- For directed graphs, $D_{\to, i} = \sum_{j \neq i} X_{ji}$ is the incoming degree of node $i$ (resp. $D_{\leftarrow, i} = \sum_{j \neq i} X_{ij}$ outcoming degree).
Erdős Rényi random graph

Erdős Rényi model

Undirected graph with no self-loops, where \( \{X_{ij}\}_{1 \leq i < j \leq n} \) are i.i.d. with distribution \( \mathcal{B}(p) \).

Characteristics

- Formulated by Erdős and Rényi in the late 50’s,
- Huge literature, describing phase transitions behaviors as \( n \to \infty \) and \( p \to 0 \) (existence of a giant component).
- Many links with branching processes.

R. Durrett.

Drawbacks

- Independence and identical distribution hypothesis both are not realistic.
- The degree distribution is \( \mathcal{B}(n, p) \approx \mathcal{P}(\lambda) \) where \( \lambda = np \) and thus does not follow a power law.
The power-law phenomenon (or scale free distribution)

- During the 00’s, many authors focused on the degree distribution of observed networks and claimed it always follows a power law
  \[ \mathbb{P}(D_i = d) = cd^{-\alpha}, \alpha \text{ being the exponent of the power law.} \]
- Some (few) nodes have a very large degree: hubs.
- They started describing networks distributions by specifying the distribution of \( \{D_i\}_{i \in V} \).
Degree distribution (power law, fixed degree . . .)

Fixed degree distribution

- Let \((d_1, \ldots, d_n)\) be the degrees of an observed graph,
- The *null* model is obtained by sampling in the set of graphs with the same degree distribution \(\sim\) rewiring algorithm.
Degree distribution (power law, fixed degree ...) 

Fixed degree distribution

- Let \((d_1, \ldots, d_n)\) be the degrees of an observed graph,
- The *null* model is obtained by sampling in the set of graphs with the same degree distribution \(\leadsto\) rewiring algorithm.
**Fixed degree distribution**

- Let \((d_1, \ldots, d_n)\) be the degrees of an **observed** graph,
- The *null* model is obtained by **sampling** in the set of graphs with the same degree distribution \(\rightsquigarrow\) **rewiring algorithm**.
Degree distribution (power law, fixed degree . . .)

Fixed degree distribution

- Let \((d_1, \ldots, d_n)\) be the degrees of an observed graph,
- The null model is obtained by sampling in the set of graphs with the same degree distribution \(\leadsto\) rewiring algorithm.

![Diagram of a graph with nodes 1, 2, \ldots, n and connections between them.

1 2 \ldots i \ldots j \ldots n

\]
Fixed degree distribution

- Let \((d_1, \ldots, d_n)\) be the degrees of an observed graph,
- The *null* model is obtained by sampling in the set of graphs with the same degree distribution \(\leadsto\) rewiring algorithm.

Note that sampling in this model is expensive. Alternative?
Degree distribution (power law, fixed degree . . .)

Mean degree distribution

- Let \( (d_1, \ldots, d_n) \) be the degrees of an observed graph and \( d_+ = \sum_i d_i \).
- Let \( \{X_{ij}\}_{1 \leq i < j \leq n} \) be independent with \( X_{ij} \sim B(p_{ij}) \) and \( p_{ij} = \frac{d_i d_j}{C} \), where \( C \) is a normalizing cst s.t. \( p_{ij} \in (0, 1) \). For instance \( C = \max_{i \neq j} d_i d_j \).
- Contrarily to 'fixed-degree' model, we do not have \( D_i = d_i \).
- Instead, \( \mathbb{E}(D_i) = d_i \frac{d_+ - d_i}{C} \). Ideally, \( d_i \) is not too large and \( C \approx d_+ \), then \( \mathbb{E}(D_i) \approx d_i \).
- If the \( d_i \)'s are not too large with respect to \( n \), then one can take \( C = C_0 := \sum_i d_i \frac{(d_+ - d_i)}{d_+} \). Then, one gets exactly \( \frac{1}{n} \sum_i \mathbb{E}(D_i) = \frac{d_+}{n} \).
Advantages and drawbacks of degree distributions

- Mean degree distribution induces independent but non i.d. edges $X_{ij} \sim \mathcal{B}(p_{ij})$. Too many parameters to be fitted to data! Mean degree fixed them to $p_{ij} \propto d_id_j$.
- Degree distribution alone does not capture all the information encoded in the graph.
Preferential attachment (dynamic) I

- Start with a small initial graph $G_0 = (V_0, E_0)$.
- At time $t$, add a new node $i_t$. For each previous node $j \in V_0 \cup \{i_1, \ldots, i_{t-1}\}$, draw edge $(i_t, j)$ with prob. $d_{j,t}/d_{+,t}$, where $d_{j,t}$ is the degree of $j$ at time $t$.

Advantages and drawbacks

- Generative model,
- Explains the power law distribution,
- Pbm of parameter choice \((V_0, E_0, t, \ldots)\).
Probabilistic models

Here, we are going to focus on (static) ‘statistical’ models,

- Exponential random graph model (ERGM).
- Stochastic block model (SBM) or MixNet.
- Overlapping stochastic block models (OSBM) or mixed membership SBM.
- Latent space models.

Some recent reviews

Exponential random graphs I

Notation

- \( X = (X_{ij})_{1 \leq i,j \leq n} \) the (binary) adjacency matrix,
- \( S(X) \) a known vector of graph statistics on \( X \)
- \( \theta \) a vector of unknown parameters

\[
P_\theta(X = x) = \frac{1}{c(\theta)} \exp(\theta^\top S(x)), \quad c(\theta) = \sum_{\text{graphs } y} \exp(\theta^\top S(y)).
\]

Statistics

- \( S(X) \) is a vector of sufficient statistics. It may contain number of edges, triangles, \( k \)-stars, \ldots and also covariates.
- Note that \( c(\theta) \) is not computable.
- Example: If \( S(x) = (x_{ij})_{1 \leq i,j \leq n} \) then
  \[
P_\theta(X = x) \propto \exp(\sum_{i,j} \theta_{ij} x_{ij}), \ i.e. \ X_{ij} \text{ are independent non i.d.}\]
  \[
  X_{ij} \sim \mathcal{B}(p_{ij}) \text{ with } p_{ij} = \frac{\exp(\theta_{ij})}{1 + \exp(\theta_{ij})}.
  \]
Imposing the constraint $\theta_{ij} = \theta$, one recovers Erdős Rényi model: $\mathbb{P}_\theta(X = x) \propto \exp(\theta S_1(x))$, where $S_1(x) = \sum_{i,j} x_{ij}$, the total number of edges is a sufficient stat. and $\hat{p} = \frac{S_1(X)}{n(n-1)/2}$.

If $S(x) = (S_1(x), S_2(x))$ with $S_2(x) = \sum_{i,j,k} x_{ij} x_{ik}$ then the variables $X_{ij}$ are non independent.

**Markov random graph:** Let $S_k(x)$ be the number of $k$-stars and $T(x) = \sum_{i,j,k} x_{ij} x_{jk} x_{ki}$ the number of triangles. For $S = (S_1, \ldots, S_{n-1}, T)$ we get

$$\mathbb{P}_\theta(X = x) \propto \exp(\sum_{k=1}^{n-1} \theta_k S_k(x) + \theta_n T(x))$$

O. Frank & D. Strauss

In practice, use only $S = (S_1, \ldots, S_k, T)$ for $k << n - 1$. 

---

Exponential random graphs II

More examples

- Imposing the constraint $\theta_{ij} = \theta$, one recovers Erdős Rényi model: $\mathbb{P}_\theta(X = x) \propto \exp(\theta S_1(x))$, where $S_1(x) = \sum_{i,j} x_{ij}$, the total number of edges is a sufficient stat. and $\hat{p} = \frac{S_1(X)}{n(n-1)/2}$.
- If $S(x) = (S_1(x), S_2(x))$ with $S_2(x) = \sum_{i,j,k} x_{ij} x_{ik}$ then the variables $X_{ij}$ are non independent.
- **Markov random graph:** Let $S_k(x)$ be the number of $k$-stars and $T(x) = \sum_{i,j,k} x_{ij} x_{jk} x_{ki}$ the number of triangles. For $S = (S_1, \ldots, S_{n-1}, T)$ we get
  $$\mathbb{P}_\theta(X = x) \propto \exp(\sum_{k=1}^{n-1} \theta_k S_k(x) + \theta_n T(x))$$

O. Frank & D. Strauss

- In practice, use only $S = (S_1, \ldots, S_k, T)$ for $k << n - 1$. 

Exponential random graphs III

Issues on parameter estimation

- Maximum likelihood estimation is difficult
- Maximum pseudo-likelihood estimators may be used [1].
  Quality of approximation?
- MCMC approaches [Hunter et al. 11]: may be slow to converge.
- Very different values of $\theta$ can give rise to essentially the same distribution.
- [CD11] established a ‘degeneracy’ of these models, which are ‘ill-posed’.


[Hunter et al. 11] D. R. Hunter, S. M. Goodreau and M. S. Handcock
ergm.userterms: A Template Package for Extending statnet.
Stochastic block models: some motivations

- Previous models do not provide a **clustering** of the nodes,
- Erdős Rényi model is too **homogeneous**: introduce heterogeneity by using **groups** (cheaper than having a parameter $p_{ij}$ for each r.v. $X_{ij}$).
- Groups could be put on **edges**, but does not take advantage of the **graph structure**. Rather put the groups on the **nodes**.
Stochastic block model (binary graphs)

- $n = 10, Z_{5\bullet} = 1$
- $X_{12} = 1, X_{15} = 0$

Binary case

- $Q$ groups (=colors $\bullet\bullet\bullet$).
- $\{Z_i\}_{1 \leq i \leq n}$ i.i.d. vectors $Z_i = (Z_{i1}, \ldots, Z_{iQ}) \sim \mathcal{M}(1, \pi)$, where $\pi = (\pi_1, \ldots, \pi_Q)$ group proportions. $Z_i$ is not observed,
- Observations: edges indicator $X_{ij}, 1 \leq i < j \leq n$,
- Conditional on the $\{Z_i\}'s$, the random variables $X_{ij}$ are independent $\mathcal{B}(p_{Z_i Z_j})$. 
Stochastic block model (weighted graphs)

Weighted case

- Observations: weights $X_{ij}$, where $X_{ij} = 0$ or $X_{ij} \in \mathbb{R}^s \setminus \{0\}$,
- Conditional on the $\{Z_i\}$'s, the random variables $X_{ij}$ are independent with distribution

$$
\mu_{Z_iZ_j}(\cdot) = p_{Z_iZ_j}f(\cdot, \theta_{Z_iZ_j}) + (1 - p_{Z_iZ_j})\delta_0(\cdot)
$$

(Assumption: $f$ has continuous cdf at zero).
Results

- **Identifiability** of parameters [AMR09, AMR11].
- **Parameter estimation / node clustering** procedures: 
  *computation of the likelihood* is not feasible (sum over $Q^n$ terms),
  *exact EM* approach is not possible,
  instead, *variational EM* or variants.

In some cases, other specific methods may be developed 
(ex: [AM12])

- **Model selection**: ICL criteria.

Identifiability of parameters in latent structure models with many observed variables, Ann. Statist., 2009.

Parameter identifiability in a class of random graph mixture models, JSPI, 2011.

New consistent and asymptotically normal estimators for random graph mixture models, JRSSB, 2012.
Variational EM algorithm in SBM

Let $\ell_n^c(\theta) := \log P_\theta(Z_{1:n}, \{X_{ij}\}_{ij})$ be the complete log-likelihood of the model.

Why EM is not possible

- EM algorithm computes $Q(\theta, \theta') := \mathbb{E}_{\theta'}(\ell_n^c(\theta) | \{X_{ij}\}_{ij})$,
- Requires the knowledge of the distribution of $Z_{1:n}$ conditional on $\{X_{ij}\}_{ij}$
- In many setups (mixtures, HMM), this distribution factorizes: $P(Z_{1:n} | \{X_{ij}\}_{ij}) = \prod_{k=1}^{n} P(Z_k | \{X_{ij}\}_{ij})$
- This is not the case in SBM. Because of the structure of the DAG

```
\cdots Z_i Z_j Z_k \cdots
  \downarrow \quad \downarrow \quad \downarrow
X_{ik} \quad X_{ij} \quad X_{jk}
```
Variational EM algorithm in SBM

Principle of the variational EM

- Idea: Replace $P(Z_{1:n} | \{X_{ij}\}_{ij})$ by its best approximation among the factorized distributions $q(Z_{1:n}) := \prod_{k=1}^{n} q_k(Z_k)$.

- More rigorously, for any distribution $q$ on $\{1, \ldots, Q\}^n$, let
  \[
  \mathcal{L}(q, \theta) = \sum_{z_{1:n}} q(z_{1:n}) \log \frac{P_{\theta(z_{1:n}, \{X_{ij}\}_{ij})}}{q(z_{1:n})}.
  \]
  Then we have
  \[
  \log P_{\theta}(\{X_{ij}\}_{ij}) = \mathcal{L}(q, \theta) + KL(q(\cdot) || P_{\theta}(Z_{1:n} = \cdot | \{X_{ij}\}_{ij})) \geq \mathcal{L}(q, \theta).
  \]

- Minimizing $KL$ w.r.t. $q \leftrightarrow$ Maximizing the lower bound $\mathcal{L}(q, \theta)$ w.r.t. $q$.

Algorithm description

- Initialize the parameter $\theta^0$,

- Iterate:
  - E-step: $\theta$ is fixed, maximize $\mathcal{L}(q, \theta)$ w.r.t. $q$,
  - M-step: $q$ is fixed, maximize $\mathcal{L}(q, \theta)$ w.r.t. $\theta$. 
Variational EM algorithm in SBM

References


Variants

▶ Variational Bayes


▶ Online variational EM

Model selection criteria in SBM ([DPR08, LBA12])

- BIC can not be computed as the maximum likelihood is still unknown
- Replace the likelihood by another (close) quantity

Integrated classification likelihood (ICL)

When convergence of variational EM is attained (step \( K \)), fix \( \hat{\theta} := \theta^K \) and let \( \hat{Z}_i = (\hat{Z}_{i1}, \ldots, \hat{Z}_{iQ}) := (q^K_i(1), \ldots, q^K_i(Q)) \) be the estimated posterior distribution of node \( i \). Then define

\[
ICL(Q) := \log P_{\hat{\theta}}(\hat{Z}_{1:n}, \{X_{ij}\}_{ij}) - \frac{N(Q)}{2} \log n,
\]

where \( N(Q) \) is the number of parameters of SBM with \( Q \) groups. Then

\[
\hat{Q} := \text{Argmin}_Q ICL(Q).
\]

C. Biernacki, G. Celeux and G. Govaert
Other properties

- Behavior of the nodes posterior dist. / Quality of variational approx. ?
  → the groups posterior distribution converges to a Dirac mass at the true groups values

- Consistency of the MLE ?
  → the MLE of the parameter converges to the true parameter value.

A. Celisse, J.-J. Daudin and L. Pierre

Mariadassou, M. and Matias, C. Convergence of the groups posterior distribution in latent or stochastic block models, Bernoulli, to appear 2014.
Overlapping SBM / Mixed membership SBM

Figure: Overlapping mixture model. Source: Palla et al., Nature, 2005.

Nodes may belong to many classes.


OSBM [Latouche et al. 11a]

Model

- $Z_i = (Z_{i1}, \ldots, Z_{iQ}) \sim \prod_{q=1}^{Q} \mathcal{B}(\pi_q)$
- $X_{ij} | Z_i, Z_j \sim \mathcal{B}(g(p_{Z_iZ_j}))$ where $g(x) = (1 + e^{-x})^{-1}$ (logistic function) and

$$p_{Z_iZ_j} = Z_i^T W Z_j + Z_i^T U + V^T Z_j + \omega$$

$W$ is a $Q \times Q$ real matrix while $U$ and $V$ are $Q$-dimensional real vectors and $\omega$ real number.

Results [Latouche et al. 11a]

- Parameter’s identifiabilty
- Variational Bayes approach + variational logistic Bayes
- Model selection criterion

Issues

- Quality of (double) variational approximation ?
Latent space models [Handcock et al. 07]

Model

- $Z_i$ i.i.d. vectors in a latent space $\mathbb{R}^d$.
- Conditional on $\{Z_i\}$, the $\{X_{ij}\}$ are independent Bernoulli
  \[
  \text{log-odds}(X_{ij} = 1|Z_i, Z_j, U_{ij}, \theta) = \theta_0 + \theta_1^T U_{ij} - \|Z_i - Z_j\|,\]
  where log-odds($A$) = log $\mathbb{P}(A)/(1 - \mathbb{P}(A))$ ; $\{U_{ij}\}$ set of covariate vectors and $\theta$ parameters vector.
- This may be extended to weighted networks
Latent space models [Handcock et al. 07]

Results [Handcock et al. 07]

- Two-stage maximum likelihood or MCMC procedures are used to infer the model’s parameters
- Assuming $Z_i$ sampled from mixture of multivariate normal, one may obtain a clustering of the nodes.

Issues

- No model selection procedure to infer the ‘effective’ dimension $d$ of latent space and the number of groups.

Outline

Some statistical networks models
  The ‘most famous’ ones
  Exponential random graphs
  (Overlapping) Stochastic block models
  Latent space models

Analyzing networks: (probabilistic) node clustering

Applications to biological networks
Clustering the nodes of a network

Probabilistic approach

- Using either mixture or overlapping mixture models, one may recover nodes groups.
- These groups reflect a common ‘connectivity behaviour’.

Non probabilistic approach = community detection

- Many clustering methods try to group the nodes that belong to the same clique.
- Here the nodes in the same groups tend to be connected with each other.
Major difference between probabilistic/non-probabilistic approach

Observation of

may lead to either

MixNet model

Clustering based on cliques
Remaining challenges

- Dynamic clustering of networks
Outline

Some statistical networks models
  The ‘most famous’ ones
  Exponential random graphs
  (Overlapping) Stochastic block models
  Latent space models

Analyzing networks: (probabilistic) node clustering

Applications to biological networks
Transcription regulatory network (TRN) of *E. coli* [PMDCR09]

TRN description

- nodes = operon (groups of genes acting together)
- link if one operon encodes a transcription factor that directly regulates another operon

Analysis

- Clustering of the graph with SBM, using 5 groups (ICL criterion)
TRN of *E. coli* [PMDCR09]

Another operon. Such networks have been shown to share some important properties, such as a relative sparseness, a very low number of feedback circuits, and a hierarchical organization [13]. Thus grouping operons based on their connectivity structure appears essential to understand the wiring diagram of such complex networks. In this paragraph, we consider the connex component of the the *E. Coli* TRN [14].

Summarizing regulatory structure: the MixNet representation

The clustering results with 5 classes (given by the ICL criterion) gives a rough picture of the network's structure. The connectivity matrix $S$ of the TRN is characterized by (i) empty rows and (ii) small diagonal elements (Table 1):

(i) means that some groups are made of strictly regulated operons (nodes that receive edges only), and (ii) that there is no community structure, i.e. there is no group which is heavily intra-connected and poorly inter-connected. This result is coherent with the structure of regulatory circuits which form cascades of regulations without feedback [13], meaning that nodes do not share modularity patterns in this regulatory network. Figure 1 indicates that the majority of operons are regulated by very few nodes. At this resolution level, the network is summarized into regulated operons (groups 1 and 4), which receive edges only. These two groups are distinguished based on their regulatory elements: operons of group 4 are regulated by crp only (which makes its own group), whereas operons of group 1 are regulated by many cross-talking elements (group 2, 3, and 5).

Meta Motifs of regulation

It has been shown that some motifs like the popular Feed Forward Loop constituted a core structure of the *E. Coli* regulatory network [14]. When looking at Figure 1, it appears that MixNet exhibits the same global structures at the group level. Groups 5 and 4 form a Single Input Module (SIM), i.e. one TF regulating other operons that do not communicate. Similarly, groups 2-3-1 and 2-5-1 form a “meta” Feed-Forward loop. In both cases the effector group is group 1, and groups 2 and 3 can be viewed as information relays.

Getting a more detailed picture

The adaptive strategy selects 12 groups which highlight the hierarchical structure of the regulation wiring diagram (Figure 2). The majority of nodes are strictly regulated operons (groups 1, 3, 5, 8, 10), whereas regulators are clustered into small groups that are distinguished based on their connectivity patterns and on their targets. For example yhdG_fis (group 2) regulates nodes of groups 1 and 8, operons of group 9 (fnr, narL) regulate operons of group 8. MixNet can also be used to detect operons that act as global TF from the connectivity point of view. For instance, rpo operons are clustered in “regulatory” classes (operon rpoE_rseABC forms group 7 on its own). This result is not surprising though, as rpo operons are involved in the V unit of the RNA polymerase. More generally, beyond groups that are made of unique major regulatory elements, MixNet gather “regulatory-like” elements together. For instance, group 4 is made of both global TF and V factors (Table 2).

Table 1: Connectivity matrix for *E. Coli* TRN with 5 classes. The probabilities of connexion are given in percentage, and probabilities lower than 1% are not displayed.

<table>
<thead>
<tr>
<th>MixNet Classes</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>2</td>
<td>.</td>
<td>4</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>3</td>
<td>1.21</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>4</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>5</td>
<td>8.64</td>
<td>17.65</td>
<td>.</td>
<td>72.87</td>
<td>11.01</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>65.49</td>
<td>5.18</td>
<td>7.92</td>
<td>21.10</td>
<td>0.30</td>
</tr>
</tbody>
</table>
Summary graph structure indicates that the majority of operons are regulated by very few nodes: At this resolution level, the network is summarized into regulated operons (groups 1 and 4), which receive edges only. These two groups are distinguished based on their regulatory elements: operons of group 4 are regulated by crp only (which makes its own group), whereas operons of group 1 are regulated by many cross-talking elements (group 2, 3, and 5).
TRN of *E. coli* [PMDCR09]

Estimated connectivity matrix

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6.40</td>
<td>1.50</td>
<td>1.34</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.21</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>8.64</td>
<td>17.65</td>
<td>.</td>
<td>72.87</td>
</tr>
<tr>
<td></td>
<td>alpha</td>
<td>65.49</td>
<td>5.18</td>
<td>7.92</td>
<td>21.10</td>
</tr>
</tbody>
</table>

- empty rows: some groups are made of strictly regulated operons (nodes that receive edges only),
- small diagonal elements: no community structure.