1 - Models with latent variables in ecology

S. Robin

INRAE / AgroParisTech / univ. Paris-Saclay
Muséum National d'Histoire Naturelle

Winter School on Mathematical Statistics, Luxembourg, Dec'20

Outline

Part 1

Outline

Community ecology

A community is a group [...] of populations of [...] different species occupying the same geographical area at the same time. [Wikipedia]

Community ecology [...] is the study of the interactions between species in communities [...]. [Wikipedia]

Community ecology

A community is a group [. . .] of populations of [. . .] different species occupying the same geographical area at the same time. [Wikipedia]

Community ecology [. . .] is the study of the interactions between species in communities [. . .]. [Wikipedia]

Need for statistical models to

- decipher / describe / evaluate
  abiotic interactions: environmental effects on species
  biotic interactions: between-species interactions

$\rightarrow$ joint species distribution models

Community ecology

A community is a group [...] of populations of [...] different species occupying the same geographical area at the same time. [Wikipedia]

Community ecology [...] is the study of the interactions between species in communities [...]. [Wikipedia]

Need for statistical models to

- decipher / describe / evaluate
  abiotic interactions: environmental effects on species
  biotic interactions: between-species interactions

  $\rightarrow$ joint species distribution models

- describe / understand the organisation of species interaction networks

  $\rightarrow$ network models

Outline

Species abundance data

Fish species in Barents sea [FNA06]:

- 89 sites (stations)
- 30 fish species
- 4 covariates

Species abundance data

Fish species in Barents sea [FNA06]:

- 89 sites (stations)
- 30 fish species
- 4 covariates

Data:

- Y_{ij} = abundance of species j in site i
- x_i = vector of covariates for site i

Abundance table:

Hi.pl	*An.lu*	*Me.ae*	...
31	0	108	
4	0	110	
27	0	788	
13	0	295	
23	0	13	
20	0	97	
.	.	.	

Environmental covariates:

Lat.	Long.	Depth	Temp.
71.10	22.43	349	3.95
71.32	23.68	382	3.75
71.60	24.90	294	3.45
71.27	25.88	304	3.65
71.52	28.12	384	3.35
71.48	29.10	344	3.65
.	.	.	.

Species abundance data

Fish species in Barents sea [FNA06]:

- 89 sites (stations)
- 30 fish species
- 4 covariates

Data:

- Y_{ij} = abundance of species j in site i
- x_i = vector of covariates for site i

Questions:

- Do environmental conditions affect species abundances? (abiotic)
- Do species abundances vary independently? (biotic)

Abundance table:

Hi.pl	*An.lu*	*Me.ae*	...
31	0	108	
4	0	110	
27	0	788	
13	0	295	
23	0	13	
20	0	97	
.	.	.	

Environmental covariates:

Lat.	Long.	Depth	Temp.
71.10	22.43	349	3.95
71.32	23.68	382	3.75
71.60	24.90	294	3.45
71.27	25.88	304	3.65
71.52	28.12	384	3.35
71.48	29.10	344	3.65
.	.	.	.

Modelling species abundance

Multivariate count distributions:

- Gaussian models do not fit
- Not that many models for count data without restriction on the dependency [IYAR17]
- Many joint species distribution models (JSDM) resort to a latent layer [WBO+15,OTD+17,PHW18]

Modelling species abundance

Multivariate count distributions:

- Gaussian models do not fit
- Not that many models for count data without restriction on the dependency [IYAR17]
- Many joint species distribution models (JSDM) resort to a latent layer [WBO+15,OTD+17,PHW18]

Poisson log-normal (PLN) model: [AH89]

Modelling species abundance

Multivariate count distributions:

- Gaussian models do not fit
- Not that many models for count data without restriction on the dependency [IYAR17]
- Many joint species distribution models (JSDM) resort to a latent layer [WBO+15,OTD+17,PHW18]

Poisson log-normal (PLN) model: [AH89]

- Z_i = latent vector associated with site i: $\{Z_i\}_{1 \leq i \leq n}$ iid

$$Z_i \sim \mathcal{N}_p(0, \Sigma)$$

Modelling species abundance

Multivariate count distributions:

- Gaussian models do not fit
- Not that many models for count data without restriction on the dependency [IYAR17]
- Many joint species distribution models (JSDM) resort to a latent layer [WBO+15,OTD+17,PHW18]

Poisson log-normal (PLN) model: [AH89]

- Z_i = latent vector associated with site i: $\{Z_i\}_{1 \leq i \leq n}$ iid

$$Z_i \sim \mathcal{N}_p(0, \Sigma)$$

- Y_{ij} = observed abundance for species j in site i : $\{Y_{ij}\}_{1 \leq i \leq n, 1 \leq j \leq p}$ independent $| \{Z_i\}$

$$Y_{ij} \sim \mathcal{P}(\exp(x_i^\intercal \beta_j + Z_{ij}))$$

Modelling species abundance

Multivariate count distributions:

- Gaussian models do not fit
- Not that many models for count data without restriction on the dependency [IYAR17]
- Many joint species distribution models (JSDM) resort to a latent layer [WBO+15,OTD+17,PHW18]

Poisson log-normal (PLN) model: [AH89]

- Z_i = latent vector associated with site i: $\{Z_i\}_{1 \leq i \leq n}$ iid

$$Z_i \sim \mathcal{N}_p(0, \Sigma)$$

- Y_{ij} = observed abundance for species j in site i : $\{Y_{ij}\}_{1 \leq i \leq n, 1 \leq j \leq p}$ independent $\mid \{Z_i\}$

$$Y_{ij} \sim \mathcal{P}(\exp(x_i^\intercal \beta_j + Z_{ij}))$$

- Unknown parameters

$$\theta = (\beta, \Sigma)$$

(Directed) Graphical model

Definition: $p(U_1, \dots U_k)$ factorizes according to the directed acyclic graph $G = ([k], E)$ iff

$$p(U_1, \dots U_k) = \prod_{h=1}^{k} p(U_h \mid U_{\text{parent}_G(h)})$$

[Lau96] (not necessarily unique factorisation)

[1] Only random variables appears as nodes, covariates in X are considered as fixed

(Directed) Graphical model

Definition: $p(U_1, \dots U_k)$ factorizes according to the directed acyclic graph $G = ([k], E)$ iff

$$p(U_1, \dots U_k) = \prod_{h=1}^{k} p(U_h \mid U_{\text{parent}_G(h)})$$

[Lau96] (not necessarily unique factorisation)

Graphical model for PLN: Independent sites + conditionally independent abundances[1]

$$p_\theta\left(\{Z_i\}, \{Y_{ij}\}\right) = \prod_i p_\Sigma(Z_i) \times \prod_j p_\beta(Y_{ij} \mid Z_i; x_i)$$

[1] Only random variables appears as nodes, covariates in X are considered as fixed

(Directed) Graphical model

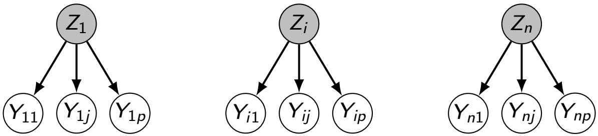
Definition: $p(U_1, \dots U_k)$ factorizes according to the directed acyclic graph $G = ([k], E)$ iff

$$p(U_1, \dots U_k) = \prod_{h=1}^{k} p(U_h \mid U_{\text{parent}_G(h)})$$

[Lau96] (not necessarily unique factorisation)

Graphical model for PLN: Independent sites + conditionally independent abundances[1]

$$p_\theta\left(\{Z_i\}, \{Y_{ij}\}\right) = \prod_i p_\Sigma(Z_i) \times \prod_j p_\beta(Y_{ij} \mid Z_i; x_i)$$



[1] Only random variables appears as nodes, covariates in X are considered as fixed

Barents' fishes

Data:

- $n = 89$ sites, $p = 30$ species, $d = 4$ covariates
- Abundance table: $Y = [Y_{ij}] \quad (n \times p)$
- Covariate table: $X = [x_{ik}] \quad (n \times d)$

Barents' fishes

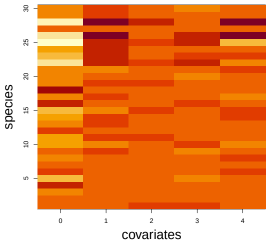
Data:

- $n = 89$ sites, $p = 30$ species, $d = 4$ covariates
- Abundance table: $Y = [Y_{ij}] \quad (n \times p)$
- Covariate table: $X = [x_{ik}] \quad (n \times d)$

Interpretation:

- β = regression coefficients
  $\rightarrow$ abiotic effects

Regression coefficients $\widehat{\beta}$: abiotic



Barents' fishes

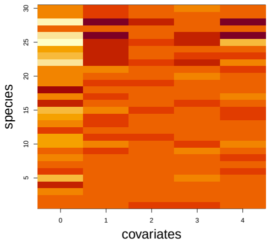
Data:

- $n = 89$ sites, $p = 30$ species, $d = 4$ covariates
- Abundance table: $Y = [Y_{ij}] \quad (n \times p)$
- Covariate table: $X = [x_{ik}] \quad (n \times d)$

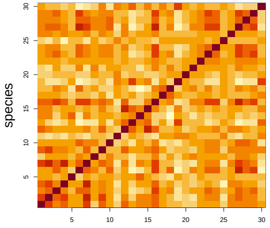
Interpretation:

- $\beta =$ regression coefficients
  $\rightarrow$ abiotic effects
- $\Sigma =$ variance of the latent layer
  $\rightarrow$ biotic associations

Regression coefficients $\widehat{\beta}$: abiotic



Covariance matrix $\widehat{\Sigma}$: biotic



Some properties of the Poisson log-normal distribution

Denoting $\Sigma = [\sigma_{jk}]$,

- Expectation (prediction):

$$\mathbb{E}(Y_{ij}) = \exp(x_i^\intercal \beta_j + \sigma_{jj}/2) =: \mu_{ij}$$

Some properties of the Poisson log-normal distribution

Denoting $\mathbf{\Sigma} = [\sigma_{jk}]$,

- Expectation (prediction):

$$\mathbb{E}(Y_{ij}) = \exp(x_i^\intercal \beta_j + \sigma_{jj}/2) =: \mu_{ij}$$

- Variance ('over-dispersion'):

$$\mathbb{V}(Y_{ij}) = \mu_{ij} + \mu_{ij}^2 (e^{\sigma_{jj}} - 1) \quad > \quad \mu_{ij}$$

Some properties of the Poisson log-normal distribution

Denoting $\Sigma = [\sigma_{jk}]$,

- Expectation (prediction):

$$\mathbb{E}(Y_{ij}) = \exp(x_i^\intercal \beta_j + \sigma_{jj}/2) =: \mu_{ij}$$

- Variance ('over-dispersion'):

$$\mathbb{V}(Y_{ij}) = \mu_{ij} + \mu_{ij}^2(e^{\sigma_{jj}} - 1) \quad > \quad \mu_{ij}$$

- Covariance:

$$\mathbb{C}\text{ov}(Y_{ij}, Y_{ik}) = \mu_{ij}\mu_{ik}(e^{\sigma_{jk}} - 1)$$

$\rightarrow$ signs are preserved:

$$\text{sign}(\sigma_{jk}) = \text{sign}(\mathbb{C}\text{ov}(Y_{ij}, Y_{ik}))$$

Outline

Models with latent variables in (community) ecology

Joint species distribution models

Models for ecological networks

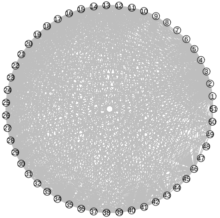
Latent-variable models

Species networks

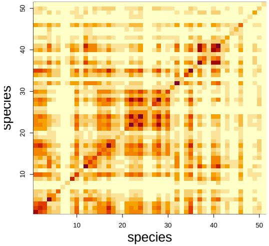
Tree network [VPDL08]:

- $n = 51$ tree species
- Y_{ij} = number of fungal parasites shared by species i and j
- x_{ij} = vector of covariates between species i and j (taxonomic, geographic, genetic distance)

Network (weighted):



Adjacency matrix (counts):



Species networks

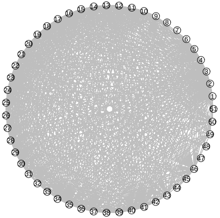
Tree network [VPDL08]:

- $n = 51$ tree species
- Y_{ij} = number of fungal parasites shared by species i and j
- x_{ij} = vector of covariates between species i and j (taxonomic, geographic, genetic distance)

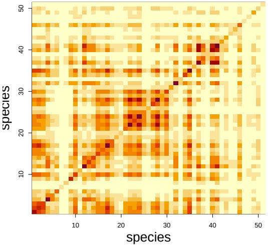
Questions:

- Is the network 'organized' in some way?
- Do the covariates contribute to explain the existence or intensity of the links?

Network (weighted):



Adjacency matrix (counts):



Modelling ecological networks

Models for (weighted) random graphs:

- Need to model the joint distribution $p(\{Y_{ij}\})$ accounting for the network structure
- Latent variable models enable to induce a row-column structure [MR14]

Modelling ecological networks

Models for (weighted) random graphs:

- Need to model the joint distribution $p(\{Y_{ij}\})$ accounting for the network structure
- Latent variable models enable to induce a row-column structure [MR14]

Stochastic block-model (SBM): [HL79,NS01]

Modelling ecological networks

Models for (weighted) random graphs:

- Need to model the joint distribution $p(\{Y_{ij}\})$ accounting for the network structure
- Latent variable models enable to induce a row-column structure [MR14]

Stochastic block-model (SBM): [HL79,NS01]

- Z_i = latent class to which node i belongs: $\{Z_i\}_{1 \leq i \leq n}$ iid

$$Z_i \sim \mathcal{M}(1, \pi = (\pi_1, \dots, \pi_K))$$

Modelling ecological networks

Models for (weighted) random graphs:

- Need to model the joint distribution $p(\{Y_{ij}\})$ accounting for the network structure
- Latent variable models enable to induce a row-column structure [MR14]

Stochastic block-model (SBM): [HL79,NS01]

- Z_i = latent class to which node i belongs: $\{Z_i\}_{1 \leq i \leq n}$ iid

$$Z_i \sim \mathcal{M}(1, \pi = (\pi_1, \dots, \pi_K))$$

- Y_{ij} = observed value (weight) of edge (i, j): $\{Y_{ij}\}_{1 \leq i,j \leq n}$ independent $| \{Z_i\}$

$$Y_{ij} \sim \mathcal{P}(\exp(x_{ij}^\intercal \beta + \alpha_{Z_i Z_j}))$$

Modelling ecological networks

Models for (weighted) random graphs:

- Need to model the joint distribution $p(\{Y_{ij}\})$ accounting for the network structure
- Latent variable models enable to induce a row-column structure [MR14]

Stochastic block-model (SBM): [HL79,NS01]

- Z_i = latent class to which node i belongs: $\{Z_i\}_{1 \leq i \leq n}$ iid

$$Z_i \sim \mathcal{M}(1, \pi = (\pi_1, \dots, \pi_K))$$

- Y_{ij} = observed value (weight) of edge (i, j): $\{Y_{ij}\}_{1 \leq i,j \leq n}$ independent $| \{Z_i\}$

$$Y_{ij} \sim \mathcal{P}(\exp(x_{ij}^\intercal \beta + \alpha_{Z_i Z_j}))$$

- Unknown parameters

$$\theta = (\pi, \beta, \alpha) \qquad + K$$

Directed graphical model

Graphical model for SBM: Independent clusters + conditionally independent edges[2]

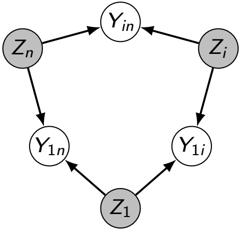
$$p_\theta\left(\{Z_i\}, \{Y_{ij}\}\right) = \prod_i p_\pi(Z_i) \times \prod_{i,j} p_{\alpha,\beta}(Y_{ij} \mid Z_i, Z_j; x_{ij})$$

[2]again: no node for the covariates in X, which are fixed

Directed graphical model

Graphical model for SBM: Independent clusters + conditionally independent edges[2]

$$p_\theta\left(\{Z_i\}, \{Y_{ij}\}\right) = \prod_i p_\pi(Z_i) \times \prod_{i,j} p_{\alpha,\beta}(Y_{ij} \mid Z_i, Z_j; x_{ij})$$



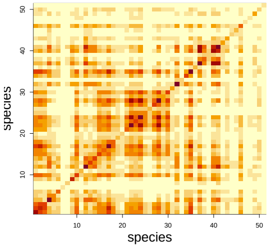
[2]again: no node for the covariates in X, which are fixed

SBM for the tree network

Data:

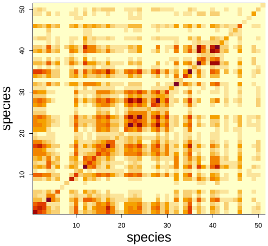
- $n = 51$ species
- Weighted adjacency matrix: $Y = [Y_{ij}]$
- No covariate (for the time being)

Observed adjacency matrix:



SBM for the tree network

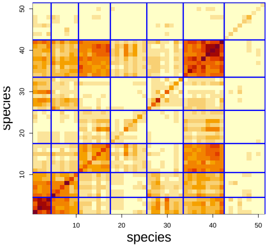
Observed adjacency matrix:



Data:

- $n = 51$ species
- Weighted adjacency matrix: $Y = [Y_{ij}]$
- No covariate (for the time being)

Clustered matrix:



Interpretation:

- π = group proportions
- α = matrix of between-groups intensities

Many types of block-models

Emission distribution: Edges can be

- Binary (presence/absence): Bernoulli
- Weighted: normal, Poisson, ...
- Multivariate (multiplex): multivariate normal, mixed multivariate distribution
- Dynamic (see Part 3)

Many types of block-models

Emission distribution: Edges can be

- Binary (presence/absence): Bernoulli
- Weighted: normal, Poisson, ...
- Multivariate (multiplex): multivariate normal, mixed multivariate distribution
- Dynamic (see Part 3)

Node structure:

- One type of nodes: symmetric or asymmetric SBM
- Two types of nodes: bipartite (see next)
- Several types of nodes: multi-layer network [BHBD19]

Bipartite networks

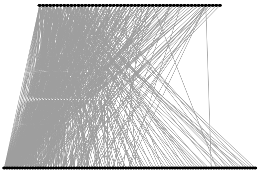
Antagonist network [BdOAN+13]:
host × parasites interactions:

- 98 hosts (fish species)
- 52 parasites

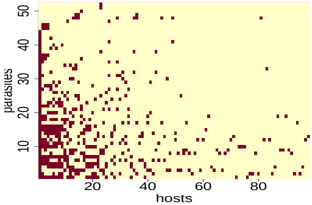
Question:
Specialized interactions ?

That is (?)
Could we determine groups of hosts and parasites that preferentially interacts (or avoid to interact)?

Network:



Adjacency matrix:



A block-model for bipartite networks

Latent block-model (LBM): [GN05]

A block-model for bipartite networks

Latent block-model (LBM): [GN05]

- Z_i = latent class to which row node i belongs: $\{Z_i\}_{1 \leq i \leq n}$ iid

$$Z_i \sim \mathcal{M}(1, \pi = (\pi_1, \dots, \pi_K))$$

A block-model for bipartite networks

Latent block-model (LBM): [GN05]

- Z_i = latent class to which row node i belongs: $\{Z_i\}_{1 \leq i \leq n}$ iid
$$Z_i \sim \mathcal{M}(1, \pi = (\pi_1, \dots, \pi_K))$$
- W_j = latent class to which row node j belongs: $\{Z_j\}_{1 \leq j \leq m}$ iid
$$W_j \sim \mathcal{M}(1, \rho = (\rho_1, \dots, \rho_L))$$

A block-model for bipartite networks

Latent block-model (LBM): [GN05]

- Z_i = latent class to which row node i belongs: $\{Z_i\}_{1 \leq i \leq n}$ iid

$$Z_i \sim \mathcal{M}(1, \pi = (\pi_1, \dots, \pi_K))$$

- W_j = latent class to which row node j belongs: $\{Z_j\}_{1 \leq j \leq m}$ iid

$$W_j \sim \mathcal{M}(1, \rho = (\rho_1, \dots, \rho_L))$$

- Y_{ij} = observed value (weight) of edge (i, j): $\{Y_{ij}\}_{1 \leq i,j \leq n}$ independent $\mid \{Z_i\}, \{W_j\}$

$$Y_{ij} \sim \mathcal{B}(\gamma_{Z_i W_j}))$$

A block-model for bipartite networks

Latent block-model (LBM): [GN05]

- Z_i = latent class to which row node i belongs: $\{Z_i\}_{1 \leq i \leq n}$ iid
$$Z_i \sim \mathcal{M}(1, \pi = (\pi_1, \dots, \pi_K))$$
- W_j = latent class to which row node j belongs: $\{Z_j\}_{1 \leq j \leq m}$ iid
$$W_j \sim \mathcal{M}(1, \rho = (\rho_1, \dots, \rho_L))$$
- Y_{ij} = observed value (weight) of edge (i, j): $\{Y_{ij}\}_{1 \leq i, j \leq n}$ independent $\mid \{Z_i\}, \{W_j\}$
$$Y_{ij} \sim \mathcal{B}(\gamma_{Z_i W_j}))$$
- Unknown parameters
$$\theta = (\pi, \rho, \gamma) + (K, L)$$

Graphical model

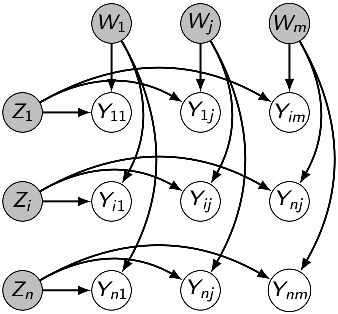
Graphical model for SBM: Independent clusters + conditionally independent edges

$$p_\theta\left(\{Z_i\}, \{W_j\}, \{Y_{ij}\}\right) = \prod_i p_\pi(Z_i) \times \prod_j p_\rho(W_j) \times \prod_{i,j} p_\gamma(Y_{ij} \mid Z_i, W_j)$$

Graphical model

Graphical model for SBM: Independent clusters + conditionally independent edges

$$p_\theta\left(\{Z_i\},\{W_j\},\{Y_{ij}\}\right) = \prod_i p_\pi(Z_i) \times \prod_j p_\rho(W_j) \times \prod_{i,j} p_\gamma(Y_{ij} \mid Z_i, W_j)$$



Latent block-model for the host-parasite network

Antagonist network [BdOAN+13]:
host × parasites interactions:

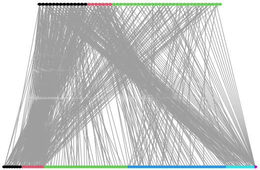
- $m = 98$ hosts (fish species)
- $n = 52$ parasites
- Adjacency matrix: $Y = [Y_{ij}]$
- No covariate

Latent block-model for the host-parasite network

Antagonist network [BdOAN+13]:
host × parasites interactions:

- $m = 98$ hosts (fish species)
- $n = 52$ parasites
- Adjacency matrix: $Y = [Y_{ij}]$
- No covariate

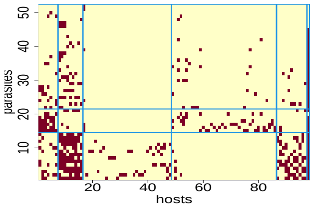
Network:



Interpretation:

- π proportions of the parasite groups
- ρ proportions of the host groups
- γ connectivities between groups of hosts and parasites

Adjacency matrix:



Outline

Two latent variable models

	Species distribution (PLN)	Network structure (SBM)
Observed (Y)	species abundances: Y_{ij} = abundance of species j in site i	species network: Y_{ij} = link intensity between species i and species j
Covariates (X)	environnemental conditions: x_i species traits x'_j	similarities between species: x_{ij} species traits x'_i
Latent (Z)	latent 'position': Z_{ij} latent variable for species j in site i	group membership: Z_i = group of species i
Parameters (θ)	latent covariance (Σ) regression coefficients (β)	group proportions (π) interactions (α) regression coefficients (β)

Some differences

Nature of the latent variables:

- **PLN** continuous, multivariate: $Z_i \sim \mathcal{N}$
- **SBM** univariate, discrete: $Z_i \sim \mathcal{M}$

Some differences

Nature of the latent variables:

PLN continuous, multivariate: $Z_i \sim \mathcal{N}$

SBM univariate, discrete: $Z_i \sim \mathcal{M}$

Modeling function of the latent variables:

PLN instrumental: comfortable multivariate Gaussian setting

$\rightarrow$ no precise ecological interpretation of the Z_i

$\rightarrow$ focus on the inference of the parameter $\theta = (\beta, \Sigma)$

Some differences

Nature of the latent variables:

PLN continuous, multivariate: $Z_i \sim \mathcal{N}$

SBM univariate, discrete: $Z_i \sim \mathcal{M}$

Modeling function of the latent variables:

PLN instrumental: comfortable multivariate Gaussian setting
- $\rightarrow$ no precise ecological interpretation of the Z_i
- $\rightarrow$ focus on the inference of the parameter $\theta = (\beta, \Sigma)$

SBM 'mechanistic': hopefully, species groups have an ecological meaning
- $\rightarrow$ easier interpretation of the Z_i
- $\rightarrow$ focus on unsupervised classification

Some differences

Nature of the latent variables:

PLN continuous, multivariate: $Z_i \sim \mathcal{N}$

SBM univariate, discrete: $Z_i \sim \mathcal{M}$

Modeling function of the latent variables:

PLN instrumental: comfortable multivariate Gaussian setting
- $\rightarrow$ no precise ecological interpretation of the Z_i
- $\rightarrow$ focus on the inference of the parameter $\theta = (\beta, \Sigma)$

SBM 'mechanistic': hopefully, species groups have an ecological meaning
- $\rightarrow$ easier interpretation of the Z_i
- $\rightarrow$ focus on unsupervised classification

Emission distribution: mostly Poisson in these lectures (a lot of count data in ecology)

$\rightarrow$ Poisson log-normal and block-models are amendable to any (parametric) distribution

References I

J. Aitchison and C.H Ho. The multivariate Poisson-log normal distribution. *Biometrika*, 76(4):643–653, 1989.

S. Bellay, E. F. de Oliveira, M. Almeida-Neto, R. M. Junior, D. P. L.and Takemoto, and J. L. Luque. Developmental stage of parasites influences the structure of fish-parasite networks. *PloS one*, 8(10):e75710, 2013.

A. Bar-Hen, P. Barbillon, and S. Donnet. Block models for multipartite networks. applications in ecology and ethnobiology. Technical Report 1807.10138, arXiv, 2019. *to appear in Statistical Modelling*.

M. Fossheim, E. M Nilssen, and M. Aschan. Fish assemblages in the Barents Sea. *Marine Biology Research*, 2(4):260–269, 2006.

G. Govaert and M. Nadif. An EM algorithm for the block mixture model. *IEEE Trans. Pattern Anal. Machine Intel.*, 27(4):643–7, 2005.

P.W. Holland and S. Leinhardt. Structural sociometry. *Perspectives on Social Network Research*, pages 63–83, 1979.

D. I Inouye, E. Yang, G. I Allen, and P. Ravikumar. A review of multivariate distributions for count data derived from the Poisson distribution. *Computational Statistics*, 9(3), 2017.

S. L. Lauritzen. *Graphical Models*. Oxford Statistical Science Series. Clarendon Press, 1996.

C. Matias and S. Robin. Modeling heterogeneity in random graphs through latent space models: a selective review. *ESAIM: Proc.*, 47:55–74, 2014.

K. Nowicki and T.A.B. Snijders. Estimation and prediction for stochastic blockstructures. *Journal of the American Statistical Association*, 96(455):1077–1087, 2001.

O. Ovaskainen, G. Tikhonov, D. Dunson, V. Grøtan, S. Engen, B.-E. Sæther, and N. Abrego. How are species interactions structured in species-rich communities? a new method for analysing time-series data. *Proceedings of the Royal Society B: Biological Sciences*, 284(1855):20170768, 2017.

G. C Popovic, F. KC Hui, and D. I Warton. A general algorithm for covariance modeling of discrete data. *Journal of Multivariate Analysis*, 165:86–100, 2018.

References II

C. Vacher, D. Piou, and M.-L. Desprez-Loustau. Architecture of an antagonistic tree/fungus network: The asymmetric influence of past evolutionary history. *PLoS ONE*, 3(3):1740, 2008.

D. I. Warton, F. G. Blanchet, R. B. O'Hara, O. Ovaskainen, S. Taskinen, S. C Walker, and F. KC. Hui. So many variables: joint modeling in community ecology. *Trends in Ecology & Evolution*, 30(12):766–779, 2015.