

Package ‘spaceNet’

May 20, 2019

Version 1.2

Date 2019-05-19

Title Latent Space Models for Multidimensional Networks

Description Latent space models for multivariate networks (multiplex) estimated via MCMC algorithm. See D Angelo et al. (2018) <arXiv:1803.07166> and D Angelo et al. (2018) <arXiv:1807.03874>.

Imports MASS, mclust (>= 5.3), permute, RcppTN, Rfast, sna, vegan

Depends R (>= 3.0)

License GPL (>= 2)

Encoding UTF-8

ByteCompile true

LazyLoad yes

NeedsCompilation no

Author Silvia D'Angelo [aut, cre],
Michael Fop [aut] (<<https://orcid.org/0000-0003-3936-2757>>),
Marco Alfò [ctb],
Thomas Brendan Murphy [ctb]

Maintainer Silvia D'Angelo <silvia.dangelo@ucd.ie>

Repository CRAN

Date/Publication 2019-05-19 22:30:03 UTC

R topics documented:

alphaRef	2
heuSearch	3
multiNet	4
spaceNet	9
vickers	10

Index	11
--------------	-----------

alphaRef

*Compute Intercept Value in the Reference Network***Description**

Computes the value for the intercept in the reference network.

Usage

```
alphaRef(Y, D = 2, sender = NULL, receiver = NULL)
```

Arguments

- Y** A three-dimensional array or list of $(n \times n)$ adjacency matrices composing the multidimensional network. A list will be converted to an array. If an array, the dimension of Y must be (n, n, K) , where n is the number of nodes and K the number of networks. Structural missing values (NA) are allowed, but no imputation is performed; see details.
- D** The dimension of the latent space, with $D > 0$. The default value is $D = 2$.
- sender, receiver** The type of node-specific sender and receiver effects to be included in the model. If specified, these effects can be set to constant ("const") or/and variable ("var"). By default, node-specific effects are not included in the model (NULL).

Details

The function computes the value for the intercept in the reference network (first network) of the multiplex. It is calculated taking into account the average effect of the latent space on edge probabilities (approximated with the constant 2) and the observed mean probability of an edge in the first network ($Y^{(1)}$):

$$p^{(1)} = \frac{\sum_{i=1}^n \sum_{j=1}^n y_{ij}^{(1)}}{n(n-1)},$$

with n the number of nodes in the network. Then, the reference intercept is computed as:

$$\log\left(\frac{p^{(1)}}{1 - p^{(1)}}\right) + 2.$$

When sender and/or receiver effects are included in the model, the intercept for the reference network is forced to be positive, see references.

Value

The function returns the intercept value in the reference network.

References

D'Angelo, S. and Murphy, T. B. and Alfò, M. (2018). Latent space modeling of multidimensional networks with application to the exchange of votes in the Eurovision Song Contest. [arXiv](#).

D'Angelo, S. and Alfò, M. and Murphy, T. B. (2018). Node-specific effects in latent space modelling of multidimensional networks. [arXiv](#).

See Also

[multiNet](#)

Examples

```
data(vickers)
alphaR <- alphaRef(vickers, D = 2) # model without node-specific effect
```

heuSearch

Heuristic model search

Description

This function performs an heuristic model search, to select which type of sender and receiver effects should be included in the model. The sender/receiver effects can be either absent, constant or variable.

Usage

```
heuSearch( Y, D = 2, nrep = 1000)
```

Arguments

Y	A three-dimensional array or list of $(n \times n)$ adjacency matrices composing the multidimensional network. A list will be converted to an array. If an array, the dimension of Y must be (n, n, K) , where n is the number of nodes and K the number of networks. Structural missing values (NA) are allowed, but no imputation is performed; see details.
D	The dimension of the latent space, with $D > 0$. The default value is $D = 2$.
nrep	The number of datasets to be simulated for each model type. By default $nrep = 1000$.

Details

The function performs an heuristic model search for the type of sender/receiver effects to be included in the model for the multidimensional network Y with n nodes and K networks.

The function simulates $nrep$ different multidimensional networks of dimensions (n, K) from the nine models arising from different combinations of sender/receiver effects: "NN", "CN", "NC", "CC", "VN", "NV", "VC", "CV", "VV". The first letter indicates the sender effect type, while the second letter denotes the receiver effect type. The effects can be constrained across the networks, with "N"

indicating no effects, "C" constant effects and "V" variable effects. The simulated multidimensional networks are then used to train a classifier, which is employed to assign the observed multiplex Y to the most plausible node-specific effect model.

Value

The function returns a list containing the following objects:

- `bestModel` The selected combination of sender and receiver effects.
- `modelProbs` The estimated probabilities that Y has been generated by the models corresponding to all the sender/receiver effects combinations.
- `modClass` The model-specific accuracies of the classifier. Rows correspond to the predicted model, while columns to the model the simulated data have been generated from.

References

D'Angelo, S. and Murphy, T. B. and Alfò, M. (2018). Latent space modeling of multidimensional networks with application to the exchange of votes in the Eurovision Song Contest. [arXiv](#).

D'Angelo, S. and Alfò, M. and Murphy, T. B. (2018). Node-specific effects in latent space modelling of multidimensional networks. [arXiv](#).

See Also

[multiNet](#)

Examples

```
data(vickers)
#heuS <- heuSearch( vickers, D = 2, nrep = 10) # nrep is set to 10 as an example.
```

multiNet

Latent Space Models for Multivariate Networks

Description

Implements latent space models for multivariate networks (multiplex) via MCMC algorithm.

Usage

```
multiNet(Y, niter = 1000, D = 2,
  muA = 0, tauA = NULL, nuA = 3,
  muB = 0, tauB = NULL, nuB = 3,
  muL = 0, tauL = NULL, nuL = 3,
  alphaRef = NULL,
  sender = c("const", "var"),
  receiver = c("const", "var"),
  covariates = NULL,
```

```
DIC = FALSE, WAIC = FALSE,
burnIn = round(niter*0.3),
trace = TRUE,
allChains = FALSE,
refSpace = NULL)
```

Arguments

Y	A three-dimensional array or list of $(n \times n)$ adjacency matrices composing the multidimensional network. A list will be converted to an array. If an array, the dimension of Y must be (n, n, K) , where n is the number of nodes and K the number of networks. Missing values (NA) are allowed; see details.
niter	The number of MCMC iterations. The default value is <code>niter = 1000</code> .
D	The dimension of the latent space, with $D > 1$. The default value is <code>D = 2</code> .
muA, muB, muL	Mean hyperparameters, see details.
tauA, tauB, tauL	Mean hyperparameters, see details.
nuA, nuB, nuL	Variance hyperparameters, see details.
alphaRef	The value for the intercept in the first network (the reference network). This value can be specified by the user on the basis of prior knowledge. By default is computed using the function <code>alphaRef</code> , see details.
sender, receiver	The type of node-specific sender and receiver effects to be included in the model. If specified, these effects can be set to constant ("const") or/and variable ("var"). By default, node-specific effects are not included in the model (NULL).
covariates	An array or a list with edge-covariates matrices. A list is automatically converted to an array. Covariates can be either continuous or discrete and must be constant throughout the views of the multiplex. The dimension of covariates is (n, n, L) , where n is the number of nodes and L the number of covariates, that is, the number of covariates matrices. Missing values (NA) are not allowed.
DIC	A logical value indicating whether the DIC (Deviance Information Criterion) should be computed. The default is <code>DIC = FALSE</code> .
WAIC	A logical value indicating whether the WAIC (Widely Available Information Criterion) should be computed. The default is <code>WAIC = FALSE</code> .
burnIn	A numerical value, the number of iterations of the chain to be discarded when computing the posterior estimates. The default value is <code>burnIn = round(niter*0.3)</code> .
trace	A logical value indicating if a progress bar should be printed.
allChains	A logical value indicating if the full parameter chains should also be returned in output. The default value is <code>allChains = FALSE</code> .
refSpace	Optional. A matrix containing a set of reference values for the latent coordinates of the nodes. Its dimension must be (n, D) , where n is the number of nodes and D the number of dimensions of the latent space. The coordinates stored in the matrix <code>refSpace</code> are compared with the estimated ones at each iteration via Procrustes correlation. High values of the correlation index indicate that the estimated coordinates are a translation and/or a rotation of the coordinates in <code>refSpace</code> .

Details

The function estimates a latent space model for multidimensional networks (multiplex) via MCMC. The model assumes that the probability of observing an arc between any two nodes is inversely related to their distance in a low-dimensional latent space. Hence, nodes close in the latent space have a higher probability of being connected across the views of the multiplex than nodes far apart. The model allows the inclusion of node-specific sender and receiver effects and edge-specific covariates.

The probability of an edge between nodes i and j in the k^{th} network is defined as:

$$P(y_{ij}^{(k)} = 1 | \Omega^{(k)}, d_{ij}, \lambda) = \frac{C_{ij}^{(k)}}{1 + C_{ij}^{(k)}}.$$

with $C_{ij}^{(k)} = \exp\{\alpha^{(k)} - \beta^{(k)}d_{ij} - \lambda x_{ij}\}$ when node-specific effects are not present and $C_{ij}^{(k)} = \exp\{\alpha^{(k)}\phi_{ij}^{(k)} - \beta^{(k)}d_{ij} - \lambda x_{ij}\}$ when they are included in the model.

The arguments of $C_{ij}^{(k)}$ are:

- The squared Euclidean distance between nodes i and j in the latent space, d_{ij}
- A coefficient λ to scale the edge-specific covariate x_{ij} . If more than one covariate is introduced in the model, their sum is considered, with each covariate being rescaled by a specific coefficient λ_l . Edge-specific covariates are assumed to be inversely related to edge probabilities, hence $\lambda \geq 0$.
- A vector of network-specific parameters, $\Omega^{(k)} = (\alpha^{(k)}, \beta^{(k)})$. These parameters are:
 - A rescaling coefficient $\beta^{(k)}$, which weights the importance of the latent space in the k^{th} network, with $\beta^{(k)} \geq 0$. In the first network (that is the reference network), the coefficient is fixed to $\beta^{(1)} = 1$ for identifiability reasons.
 - An intercept parameter $\alpha^{(k)}$, which corresponds to the largest edge probability allowed in the k^{th} network. Indeed, when $\beta^{(k)} = 0$ and when no covariate is included, the probability of having a link between a couple of nodes is that of the random graph:

$$P(y_{ij}^{(k)} = 1 | \alpha^{(k)}) = \frac{\exp\{\alpha^{(k)}\}}{1 + \exp\{\alpha^{(k)}\}}.$$

The intercepts have a lower bound corresponding to $\log\left(\frac{\log(n)}{n - \log(n)}\right)$. For identifiability reasons, the intercept of the first network needs to be fixed. Its value can be either specified by the user on the basis of prior knowledge or computed with the function [alphaRef](#).

- When node-specific effects are included in the model,

$$\phi_{ij}^{(k)} = g(\theta_i^{(k)} + \gamma_j^{(k)})$$

with :

- $\theta_i^{(k)}$ the sender effect of node i in network k .
- $\gamma_j^{(k)}$ the receiver effect of node j in network k .
- g a scalar. When both sender and receiver effects are present, $g = 0.5$; when only one type of effect is included in the model, $g = 1$.

When the sender and/or receiver effects are set to constant ("const"), each node i is assumed to have a constant effect across the different networks: $\theta_i^{(k)} = \theta_i$ and/or $\gamma_i^{(k)} = \gamma_i$. Instead, when they are set to variable ("var"), each node has a different effect across the networks: $\theta_i^{(k)}$ and/or $\gamma_i^{(k)}$.

Inference on the model parameters is carried out via a MCMC algorithm. A hierarchical framework is adopted for estimation, where the parameters of the distributions of α , β and λ are considered nuisance parameters and assumed to follow hyper-prior distributions. The parameters of these hyperpriors need to be fixed and are the following:

- tauA, tauB and tauL are the scale factors for the variances of the hyperprior distributions for the mean parameters of $\alpha^{(k)}$, $\beta^{(k)}$ and λ_l . If not specified by the user, tauA and tauB are computed as $(K - 1) / K$, if $K > 1$, otherwise they are set to 0.5. Parameter tauL is calculated as $(L - 1) / K$, if $L > 1$, otherwise it is set to 0.5.
- muA, muB and muL are the means of the hyperprior distributions for the mean parameters of $\alpha^{(k)}$, $\beta^{(k)}$ and λ_l . If not specified by the user, they are all set to 0.
- nuA, nuB and nuL are the degrees of freedom of the hyperprior distributions for the variance parameters of $\alpha^{(k)}$, $\beta^{(k)}$ and λ_l . If not specified by the user, they are all set to 3.

Missing data are considered structural and correspond to edges missing because one or more nodes are not observable in some of the networks of the multiplex. No imputation is performed, instead, the term corresponding to the missing edge is discarded in the computation of the likelihood function. For example, if either node i or j is not observable in network k , the edge (i, j) is missing and the likelihood function for network k is calculated discarding the corresponding (i, j) term. Notice that the model assumes a single common generative latent space for the whole multidimensional network. Thus, discarding the (i, j) term in the k^{th} network does not prevent from recovering the coordinates of nodes i and j in the latent space.

Value

An object of class 'multiNet' containing the following components:

n	The number of nodes in the multidimensional network.
K	The number of networks in the multidimensional network.
D	The number of dimensions of the estimated latent space.
parameters	A list with the following components: <ul style="list-style-type: none"> • alpha is a list with two components: the means of the posterior distributions and the standard deviations of the posterior distributions for the intercept parameters; • beta is a list with two components: the means of the posterior distributions and the standard deviations of the posterior distributions for the latent space coefficient parameters; • theta is a list with two components: the means of the posterior distributions and the standard deviations of the posterior distributions for the sender effect parameters; • gamma is a list with two components: the means of the posterior distributions and the standard deviations of the posterior distributions for the receiver effect parameters;

	<ul style="list-style-type: none"> • <code>lambda</code> is a list with two components: the means of the posterior distributions and the standard deviations of the posterior distributions for the covariate coefficient parameters.
<code>latPos</code>	A list with posterior estimates of means and standard deviations of the latent coordinates.
<code>accRates</code>	<p>A list with the following components:</p> <ul style="list-style-type: none"> • <code>alpha</code> is a vector with the acceptance rates for the intercept parameters; • <code>beta</code> is a vector with the acceptance rates for the latent space coefficient parameters; • <code>theta</code> is a matrix with the acceptance rates for the sender effect parameters; • <code>gamma</code> is a matrix with the acceptance rates for the receiver effect parameters; • <code>lambda</code> is a vector with the acceptance rates for the covariate coefficient parameters; • <code>latPos</code> is a vector with the acceptance rates for the latent coordinates of the nodes.
<code>DIC</code>	The Deviance Information Criterion of the estimated model. Computed only if <code>DIC = TRUE</code> in input.
<code>WAIC</code>	The Widely Available Information Criterion of the estimated model. Computed only if <code>WAIC = TRUE</code> in input.
<code>allChains</code>	<p>If <code>allChains = TRUE</code>, a list with the following components is returned:</p> <ul style="list-style-type: none"> • <code>parameters</code> is a list with the estimated posterior distributions of the model parameters: α, β, θ, γ and λ; • <code>latPos</code> is an array with the posterior distributions of the latent coordinates of each node; • <code>priorParameters</code> is a list with the estimated posterior distributions of the parameters of the prior distributions of α, β and λ.
<code>corrRefSpace</code>	A numerical vector containing the values of the Procrustes correlation between the reference space and the estimated one, computed at each mcmc iteration. Only outputted when <code>refSpace</code> is given, otherwise NULL.
<code>info</code>	<p>A list with some information on the estimated model:</p> <ul style="list-style-type: none"> • <code>call</code> contains the function call; • <code>niter</code> is the number of MCMC iterations; • <code>burnIn</code> is the number of initial iterations to discarded when computing the estimates; • <code>sender</code> is the node-specific sender effect type; • <code>receiver</code> is the node-specific receiver effect type; • <code>covariates</code> is the covariates array, if present; • <code>L</code> is the number of covariates.

References

- D'Angelo, S. and Murphy, T. B. and Alfò, M. (2018). Latent space modeling of multidimensional networks with application to the exchange of votes in the Eurovision Song Contest. [arXiv](#).
- D'Angelo, S. and Alfò, M. and Murphy, T. B. (2018). Node-specific effects in latent space modelling of multidimensional networks. [arXiv](#).

See Also[alphaRef](#)**Examples**

```
data(vickers)

it <- 10    # small number of iterations just for example

# 2-dimensional latent space model, no covariates
mod <- multiNet(vickers, niter = it, D = 2)

# 2-dimensional latent space model, sex as covariate
mod <- multiNet(vickers, niter = it, D = 2,
               covariates = sex)

# 2-dimensional latent space model, with constant sender
# effect and variable receiver effect
mod <- multiNet(vickers, niter = it, D = 2,
               sender = "const", receiver = "var")
```

spaceNet

Latent Space Models for Multivariate Networks

Description

A package for latent space models for binary multivariate networks (multiplex). The model assumes that the nodes in the multiplex lie in a low-dimensional latent space. The probability of two nodes being connected is inversely related to their distance in this latent space: nodes close in the space are more likely to be linked, while nodes that are far apart are less likely to be connected. The model allows the inclusion of node-specific sender and receiver effects and edge-specific covariates. Inference is carried out via a MCMC algorithm.

Details

The main function is `multiNet`, which estimates the latent space model via MCMC algorithm. Data can be inputted either as a list or an array. Also, edge-specific covariates in the form of a list or an array can be included in the model.

Author(s)

Silvia D'Angelo and Michael Fop.

Maintainer: Silvia D'Angelo <silvia.dangelo@ucd.ie>

References

D'Angelo, S. and Murphy, T. B. and Alfò, M. (2018). Latent space modeling of multidimensional networks with application to the exchange of votes in the Eurovision Song Contest. [arXiv](#).

D'Angelo, S. and Alfò, M. and Murphy, T. B. (2018). Node-specific effects in latent space modelling of multidimensional networks. [arXiv](#).

vickers

Vickers-Chan 7th Graders Multidimensional Network

Description

The data by Vickers et al. (1981) represent three different social relations among 29 seventh grade students in a school in Victoria, Australia. A binary covariate matrix is available and it describes if two students share the same sex or not.

Usage

```
data(vickers)
```

Details

The data by Vickers et al. (1981) are stored in a list named `vickers` and represent three different social relations among 29 seventh grade students, 12 boys and 17 girls, in a school in Victoria, Australia. Each student was asked if

- *Network 1*: he\ she got on with another given student;
- *Network 2*: he\ she considered another given student as his\ her best friend;
- *Network 3*: he\ she would have liked to work with another given student.

The matrix `sex` is an edge-specific binary covariate recording the gender of pairs of students. If a pair of students is of the same sex, the corresponding entry in the matrix is 0, otherwise is 1.

References

Vickers, M. and Chan, S. (1981). Representing Classroom Social Structure. *Melbourne: Victoria Institute of Secondary Education*.

Examples

```
data(vickers)
```

Index

`alphaRef`, [2](#), [5](#), [6](#), [9](#)

`heuSearch`, [3](#)

`multiNet`, [3](#), [4](#), [4](#), [9](#)

`print.multiNet (multiNet)`, [4](#)

`sex (vickers)`, [10](#)

`spaceNet`, [9](#)

`spaceNet-package (spaceNet)`, [9](#)

`vickers`, [10](#)