Package 'SyNet'

April 24, 2008

Type Package

Title Inference and Analysis of Sympatry Networks

Version 1.0

Date 2008-04-24

Author Daniel A. Dos Santos

Maintainer Daniel A. Dos Santos <pseudalopex_79@yahoo.com>

Depends deldir

Description Infers sympatry matrices from distributional data and analizes them in order to identify
groups of species cohesively sympatric.

License GPL

R topics documented:

Partition_Index   *Index of Network Partition*

Description

Indicates the tendency of a sympatry network to contain species aggregated into locally dense groups, despite the existence of intermediary species connecting them to a more global low density network.

Usage

```
Partition_Index(msym, n = NROW(msym))
```

Arguments

msym   Matrix object. It must be binary (1/0) and symmetric.

n   Number of nodes.

Details

The index is based on clustering coefficient measure (**C**). It considers, for each node, the maximum between its own **C** value and the lowest **C** value at its neighbourhood. The selected value is known as clustering performance after Dos Santos et al. (2008). Finally, the mean of clustering performances is obtained.

This function is for internal use.

Value

A real number between 0 and 1, where high values suggest the existence of dense groups of sympatric species.

Author(s)

Daniel A. Dos Santos

References

Dos Santos, D.A., Fernandez, H.R., Cuezzo, M.G., Dominguez, E. *Sympatry Inference and Network Analysis in Biogeography.* Systematic Biology (in press).

See Also

This statistic is used by partition.

SyNet-package   *Inference and Analysis of Sympatry Networks*

Description

Historical Biogeography focuses basically on sympatry patterns among species. **SyNet** extracts hypotheses of sympatry from available distributional evidence and integrates them into sympatry networks. **SyNet** tests the adequacy of networks to be segregated into groups of species cohesively sympatric (units of co-occurrence). It identifies units of co-occurrence from the more inclusive network where they are embedded. To do this, the program iteratively removes intermediary species that connect the units to the general network. Data from the spatial expressions of units are transferred to output files that can be opened by GIS softwares and Google Earth.

Details

Package:	SyNet
Type:	Package
Version:	1.0
Date:	2008-04-1
License:	GPL

Author(s)

Daniel A. Dos Santos

References

Dos Santos, D.A., Fernandez, H.R., Cuezzo, M.G., Dominguez, E. *Sympatry Inference and Network Analysis in Biogeogrpahy*. Systematic Biology (in press)

See Also

Sympatry inference from dot maps uses Delaunay Triangulation on the set of records considered. The package deldir is thus required to implement this action. deldir can be downloaded from the Comprehensive R Archive Network (http://cran.r-project.org/).

Examples

```
data(sciobius2x2)

#####
# Derive a sympatry matrix. >=1 for co-occurrence in at least one location.
sm <- as.matrix(sciobius2x2)%*%t(as.matrix((sciobius2x2)))
# Plot a sympatry network
forgraph <- seq(0, by = 2*pi/47, length = 47)
xcoord <- cos(forgraph)
ycoord <- sin(forgraph)
par(mfrow = c(2,2))
plot(xcoord, ycoord, main = "Basal Sympatry Network",
```

```
    sub = "Non-significant Node Ordering", xlab = "", ylab = "")
    text (xcoord*1.05, ycoord*1.05, rownames(sciobius2x2))
    for (i in 1:46)
      for (j in (i+1):47)
        if(sm[i,j] > 0) lines (c(xcoord[i], xcoord[j]), c(ycoord[i], ycoord[j]))
    # The layout is sub-optimal, and it seems chaotic. But as Saramago says:
    # 'Chaos is an order not yet decoded'.

    #####
    # Now, we order the nodes according to their NAM Status.
    namsc <- nam(gridinfer(dntable = sciobius2x2))
    codenam <- as.integer(namsc$Categories[,3])
    neworder <- order(codenam)
    plot(xcoord, ycoord, main = "Basal Sympatry Network",
    sub = "NAM Node Ordering", xlab = "", ylab = "" )
    text (xcoord*1.05, ycoord*1.05, namsc$Categories[neworder,1])
    for (i in 1:46)
      for (j in (i+1):47)
        if(sm[neworder[i],neworder[j]] > 0)
          lines (c(xcoord[i], xcoord[j]), c(ycoord[i], ycoord[j]))

    #####
    # Finally, we delete links inciding on those nodes classified
    # as isolated or intermediary ones.
    plot(xcoord, ycoord, main = "Selected Sub-network",
    sub = "Units of Co-occurrence", xlab = "", ylab = "" )
    text (xcoord*1.05, ycoord*1.05, namsc$Categories[neworder,1])
    for (i in 1:46){
      if (codenam[neworder[i]] < 1) next
      for (j in (i+1):47) {
        if (codenam[neworder[j]] < 1)
          next
        if(sm[neworder[i],neworder[j]] > 0)
          lines (c(xcoord[i], xcoord[j]), c(ycoord[i], ycoord[j]))
      }
    }
```

---

dotinfer   *Sympatry Inference from Records*

Description

Produces a sympatry matrix from available distributional data organized as a list of species records.

Usage

```
dotinfer(x)
```

Arguments

x   Object of class 'dnpoint'

Details

Species are sympatric if their ranges overlap, while they are allopatric if their ranges show spatial disjunction. Inference of sympatry can be addressed over raw distributional data (i. e., dot maps or coordinates of species records) or over distributional tables (i. e., tables of species vs. pre-defined areas that indicates occupancy or not).

In case of punctual data, sympatry is inferred by the interaction of geographical proximity and interpenetration of sets of species records. Thus, sympatry is proposed when records are close together, sharing a subjacent area of unknown boundary. A step by step explanation of the procedure can be found in the source code. In case of grid based distributional data, sympatry is inferred by co-occurrence of species in at least one pre-defined spatial unit.

The first approach uses the evidence of distribution without appealing to *a priori* delimited areas. It avoids conversion of records into ranges as a preliminar step to infer sympatry.

The second approach encodes the distribtuion of species in terms of area occupancy. It is prone to obtain results biased by the size (scale effect) and shape of spatial units considered.

Value

An object of class 'dotinference' with components:

sm   An adjacency matrix that reflects the existence (1) or not (0) of a sympatric link between species.

Label   Character vector of species label.

Author(s)

Daniel A. Dos Santos

References

Dos Santos, D.A., Fernandez, H.R., Cuezzo, M.G., Dominguez, E. *Sympatry Inference and Network Analysis in Biogeography.* Systematic Biology (in press).

See Also

The package `deldir` is required to infer sympatry from punctual records.

Examples

```
data(epiphragmophora)
dotinfer(epiphragmophora) # Infer the sympatry matrix.
```

---

epiphragmophora   *Epiphragmophora Distributional Data*

Description

Provides records of 21 species of the landsnail genus *Epiphragmophora* consisting of 145 coordinates in decimal format.

Usage

```
data(epiphragmophora)
```

Format

List of 3 elements: $Numpoints : int145$ Points :'data.frame': 145 obs. of 3 variables: $..IDsp : num[1 : 145]1111111111.....$ Longitud: num [1:145] -65.5 -65.4 -65.3 -65.2 -65.4 ... $..Latitud : num[1 : 145] - 26.8 - 26.8 - 26.7 - 26.6 - 26.8...$ Label : chr [1:21] "argentina" "cryptomphala" "escoipensis" "guevarai" ... - attr(*, "class")= chr "dnpoint"

Details

Data were previously incorporated to an object of class 'dnpoint'.

Source

M. G. Cuezzo, curator of Malacological Collection of Miguel Lillo Museum, Tucuman, Argentina.

References

Cuezzo, M. G. 2006. *Systematic revision and cladistic analysis of Epiphragmophora Doering from Argentina and Southern Bolivia (Gastropoda: Stylommatophora: Xanthonychidae).* Malacologia 49:121-188.

Dos Santos, D.A., Fernandez, H.R., Cuezzo, M.G., Dominguez, E. 2008. *Sympatry Inference and Network Analysis in Biogeography.* Systematic Biology (in press).

Examples

```
data(epiphragmophora)
namep <- nam(dotinfer(epiphragmophora))

# Displays species labels belonging to each NAM Status category
split(namep$Categories[,1], namep$Categories[,2])
```

---

gridinfer	*Sympatry Inference from Grids*

Description

Produces a sympatry matrix from distributional data organized as tables of species vs. pre-defined spatial units.

Usage

```
gridinfer(file = NULL, dntable = NULL, sp_row = TRUE)
```

Arguments

`file`	Character string naming the ASCII file to read it. The file is read by `read.table` and contains an species-by-grids matrix of presence/absence. Entries must be non-negative and scores higher than zero are interpreted as presence.
`dntable`	A matrix or data frame object with non-negative entries. It is an species-by-grids distributional table. Entries must be non-negative and scores higher than zero are interpreted as presence.
`sp_row`	Logical. If `TRUE` rows are interpreted as species and columns as grids, otherwise the opposite is considered.

Details

Species are sympatric if their ranges overlap, while they are allopatric if their ranges show spatial disjunction. Inference of sympatry can be addressed over raw distributional data (i. e., dot maps or coordinates of species records) or over distributional tables (i. e., tables of species vs. pre-defined areas that indicates occupancy or not).

In case of punctual data, sympatry is inferred by the interaction of geographical proximity and interpenetration of sets of species records. Thus, sympatry is proposed when records are close together, sharing a subjacent area of unknown boundary. A step by step explanation of the procedure can be found in the source code. In case of distributional data based on grids, sympatry is inferred by co-occurrence of species in at least one pre-defined spatial unit.

The first approach uses the evidence of distribution without appealing to *a priori* delimited areas. It avoids the conversion of records into ranges as a preliminar step to infer sympatry.

The second approach encodes the distribtuion of species in terms of area occupancy. It is prone to obtain results biased by the size (scale effect) and shape of spatial units considered.

Value

An object of class `gridinference`, which is a list with components:

`sm`	An adjacency matrix that reflects the existence (1) or not (0) of a sympatric link between species.
`Label`	Character vector of species labels.
`Dngrid`	Distributional table of species-by-grids.

Note

Given a distribuional matrix $\mathbf{D}$, where rows are species and columns grids, the sympatry matrix $\mathbf{S}$ is firstly inferred through the matrix mulplication of $\mathbf{D}$ with its transpose. Then, entries are dichotomized, becoming $\mathbf{S}$ in a binary matrix.

Author(s)

Daniel A. Dos Santos

References

Dos Santos, D.A., Fernandez, H.R., Cuezzo, M.G., Dominguez, E. *Sympatry Inference and Network Analysis in Biogeography.* Systematic Biology (in press).

See Also

Objects of class gridinference can be submitted to functions outgrid and nam.

Examples

```
data(sciobius2x2)
gridinfer(dntable = sciobius2x2)$sm # Displays the sympatry matrix
```

---

hullinfer   *Sympatry Inference from Convex Hulls*

Description

Realizes all pairwise comparisons of overlapping between species ranges encoded as convex hulls.

Usage

```
hullinfer(x)
```

Arguments

x   An object of class dnpoint.

Details

chull is used to obtain the convex hulls.
The internal analysis of geometric overlap is carried out by the function overlaphull.

Value

An object of class hullinference, which is a list with the following elements:

sm   An adjacency matrix that reflects the existence (1) or not (0) of a sympatric link
                between species.

Label   A character vector of species labels.

HullVertex   List of convex hulls by species. Each main element shows the x and y coordi-
                nates of points lying on the convex hulls, in clockwise order.

Author(s)

Daniel A. Dos Santos

Examples

```
#####
# We create an object of class 'dnpoint' that includes many degeneracies
IDsp <- c(1,1,2,2,3,3,3,4,4,4,4,5,5,5)
Longitud <- c(1,1,3,3,0,4,5.5,1,1,2,2,1.5,3.5,3.5)
Latitud <- c(5,8,0,2,1,7,0,4,9,4,6,2,2,3)
Label <- c("A", "B", "C", "D", "E")
example <- list(Numpoints = 14, Points = data.frame(IDsp = IDsp,
          Longitud = Longitud, Latitud = Latitud), Label = Label)
class(example) <- "dnpoint"
rslt <- hullinfer(example)

#####
# Draw ranges
plot(Longitud, Latitud)
for (i in 1:5)
  polygon(rslt$HullVertex[[i]], border = i)
```

---

nam	*Sympatry Network Analysis*

Description

Performs the analysis of a sympatry network oriented to detect groups of species cohesively sympatric (units of co-occurrence) via iterative removal of nodes with high betweenness score (intermediary species).

Usage

```
nam(msym)
```

Arguments

msym   Object of classes 'dotinference', 'gridinference' or 'hullinference'.

Details

The input argument `msym` contains a sympatry matrix $\mathbf{S}$ of order n (= number of species). Each entry Sij is 1 if there is a sympatric link between species i and j, while Sij is 0 if allopatry is suggested. The diagonal elements are 1 because sympatry is reflexive. Furthermore, sympatry matrix is symmetric due to reciprocal nature of this relationship.

Sympatry matrices are adjacency matrices, and finding the connection patterns in the associated network is the scope of `nam` function. In a given sympatry network, nodes represent species and edges sympatric links. `nam` is an iterative process of node removal to isolate subsets of nodes with within-group sympatry and between-group allopatry (units of co-occurrence). Units of co-occurrence are generally embedded in the global network due to intermediary nodes connecting them.

`nam` identifies and removes nodes with the highest intermediacy. Intermediacy is evaluated with the betweenness measure (Freeman, 1977). After the node removal, a sub-network is generated and the process is repeated until it arrives to a sub-network with all nodes with zero betweenness. In this way, `nam` produces a serie of sub-networks, one of them holding the units of co-occurrence.

As units of co-occurrence are being segregated, the overall clustering performance (**OCP**) increases. This index reflects the change in clustering performance along all nodes of a sub-network with respect to the basal condition (Dos Santos et al., 2008). The sub-network maximizing **OCP** is selected. Here, the components with 3 or more nodes represent the units of co-occurrence we are interested on.

Value

An object of class `nam`, which is a list with components:

`LastNet`	Integer vector indicating the last network (or sub-network) where a given node was found. Zero for the basal network and >0 for succesive sub-networks.
`Betweenness`	Highest betweenness value recorded at the respective instance of removal process.
`OCPtrajectory`	
	Serie of **OCP** values recorded along the removal process.
`Selected`	Index of sub-network selected.
`Categories`	Data frame that arranges species into four categories of **NAM** Status. See Note below.

Simultaneously, the function displays a graph showing the evolution of **OCP** throughout the removal process (y-axis: index value; x-axis: sub-network considered).

Note

Betweenness measure was calculated with Newman's algorithm (Newman, 2001).

NAM classifies species into four categories: Intermediary species, Isolated species, Diad and Unit of Co-occurrence(UC).

Author(s)

Daniel A. Dos Santos

References

Freeman, L. C. 1977. *A set of measures of centrality based on betweenness.* Sociometry 40:35-41.

Newman, M. E. J. 2001. *Scientific collaboration networks. II. Shortest paths, weighted networks and centrality.* Phys. Rev. E 64:016132.

Dos Santos, D. A., Fernandez, H. R., Cuezzo, M. G., Dominguez, E. 2008. *Sympatry Inference and Network Analysis in Biogeography.* Systematic Biology (in press).

See Also

Objects of class 'dotinference' are produced with function `dotinfer`.
Objects of class 'gridinference' are produced with function `gridinfer`.
Objects of class 'hullinference' are produced with function `hullinfer`.
Previous to **NAM** analysis, is necessary to test the network adequacy to be segregated into units of co-occurrence:`partition`.
Concept of clustering performance is summarized in Details section of `partition` documentation.

Examples

```
data(epiphragmophora)
nam(dotinfer(epiphragmophora))
```

---

outgearth	*Generation of a KML File*

Description

Produces Google Earth placemarks from punctual species records.

Usage

```
outgearth(x, dn, outfile = "namgearth.kml")
```

Arguments

x	Object x of class 'nam'.
dn	Object dn of class 'dnpoints'.
outfile	A character string naming the .kml file to write to.

Details

The Homepage of Google Earth is available at: http://earth.google.com/

Author(s)

Daniel A. Dos Santos

References

A complete documentation about KML files are available at: http://code.google.com/apis/kml/documentation/kml_tut.html

Examples

```
data(epiphragmophora)
outnam <- nam(dotinfer(epiphragmophora))
outgearth(outnam, epiphragmophora, "proof") ## Find the file in the
                                            ## working directory
```

---

outgis   *Generating Files for GIS Softwares*

Description

Generates a .txt file that can be imported by GIS softwares (e.g. DIVA-GIS) in order to convert the data in a shapefile of points.

Usage

```
outgis(x, dn, outfile = "namgis.txt")
```

Arguments

x	Objetc of class 'nam'.
dn	Object of class 'dnpoint'.
outfile	A character string naming the .txt file to write to.

Details

The generated file has four fields separated by comma: Species Label, Longitude, Latitude and NAM Status. The first line of file contains the field labels. From second line on, respective field information.

Note

DIVA-GIS is a free software widely used in biogeographical research. For DIVA-GIS users, the generated .txt file can be opened from menu DATA -> IMPORT POINTS TO SHAPEFILE -> FROM TEXT FILE (.TXT).

Author(s)

Daniel A. Dos Santos

See Also

The homepage of DIVA-GIS is at http://www.diva-gis.org/.
Function nam for generating objects of 'nam' class.
Function read.coord for reading records and generating objects of class 'dnpoint'.

Examples

```
data(epiphragmophora)
p <- dotinfer(epiphragmophora)
x <- nam(p)
outgis(x, epiphragmophora) ##See the file in the working directory
```

---

outgrid   *Spatial Expression of Grid Data*

Description

Displays a list with operative geographical units (OGUs) associated to each unit of co-occurrence.

Usage

```
outgrid(x, namx = NULL)
```

Arguments

x   Object of class 'gridinference'.

namx   Object of class 'nam'. If NULL nam function is implemented.

Value

A list of character vectors. Elements correspond to spatial expressions of units of co-occurrence detected through NAM analysis.

Author(s)

Daniel A. Dos Santos

See Also

nam for generating objects of class nam. gridinfer for generating objects of class gridinference.

Examples

```
data(sciobius2x2) # Obtain grid data about Sciobius distribution.
arg1 <- gridinfer(dntable = sciobius2x2)
arg2 <- nam(arg1)
outgrid(arg1, arg2)
```

---

overlaphull   *Intersection Between Convex Hulls*

Description

Determines if species ranges (encoded as convex hulls) overlap.

Usage

```
overlaphull(sp1, sp2)
```

Arguments

sp1	x and y coordinates of points lying on convex hull of first species. Coordinates are provided as a list.
sp2	The same, but for second species.

Details

Auxilliary function and not intended to be used by final user.

Value

One (1) for intersecting convex hulls, otherwise zero (0)

Author(s)

Daniel A. Dos Santos

See Also

function hullinfer calls this method.

---

partition	*Test of Adequacy*

Description

Tests the capability of a sympatry network to contain units of co-occurrence.

Usage

```
partition(iptsymp, replica = 1)
```

Arguments

iptsymp	Matrix object. It must be binary (1/0) and symmetric.
replica	Number of random matrices to be produced.

Details

The input argument iptsymp corresponds to any adjacency matrix that reflects the incidence (1) or not (0) of a sympatric link between pairs of species.

Random matrices are generated according to a Bernoulli model. A random number between 0 and 1 is generated for each pair of species or cell. If this number is lower than the density of the network associated to the observed matrix, a 1 score is assigned.

The statisitc corresponds to the Partition Index (**PI**). The test counts the number of random matrices that yields a **PI** value higher or equal to the observed one.

Value

If `replica` > 1, a list is generated with the following components:

`$Input`	Sympatry matrix observed.
`$ProbTie`	Density of network associated to.
`$NodeNr`	Number of nodes considered.
`$PIobserved`	Observed Partition Index.
`$PIrandomized`	
	Statistical summary of randomized Partition Index.
`$ProbTie`	Fraction of `replica` where the randomized Partition Index is >= the observed one.

If `replica` = 1, the Partition Index of the observed matrix is calculated.

Author(s)

Daniel A. Dos Santos

References

Dos Santos, D.A., Fernandez, H.R., Cuezzo, M.G., Dominguez, E. *Sympatry Inference and Network Analysis in Biogeography.* Systematic Biology (in press).

See Also

The statistic **PI** is calculated by the function `Partition_Index`

Examples

```
data(epiphragmophora)
p <- dotinfer(epiphragmophora)
partition(p[[1]], 100)
```

---

read.coord	*Read a List of Coordinates*

Description

Organizes the distributional information about species. Data are provided as set of punctual records with latitude and longitude coordinates.

Usage

```
read.coord(inputfile = "")
```

Arguments

`inputfile`	A character string naming the .txt file to read it.

Details

The input file must contain three fields separated by comma. A header is necessary to identify latitude from longitude. The first field must correspond to the ID of species. The second and third fields corresponds to longitude and latitude wihtout distinction of order priority. The rest of file corresponds to records.

Value

An object of class `dnpoint`, which is a list with elements:

`Numpoints`	Number of read points.
`Points`	Data frame with columns (1) index of species, (2) Longitude and (3) Latitude. Coordinates are in decimal format.
`Label`	Character vector of species labels.

Author(s)

Daniel A. Dos Santos

See Also

Objects of class `dnpoint` are required by the functions `dotinfer`, `hullinfer`, `outgis` and `outgearth`.

Examples

```
#####
# You can recognize the format of a typical input file
# in the following created .txt:
write(c("sp", "latitude", "longitude"), file= "proof.txt", 3, TRUE, sep = ",")
for (i in 1:10)
  write(c(LETTERS[i], i*1.5, i*-8.6),file= "proof.txt", 3, TRUE, sep = ",")
proof <- read.coord("proof.txt") # Put getwd() to identify path where
                                 # the file 'proof.txt' was located


#####
unlink("proof.txt") # Delete
```

---

sciobius2x2   *Sciobius Grid Occupancy Data*

Description

Presence/absence data of 47 South African *Sciobius* species along a square grid of 2 x 2 degrees.

Usage

```
data(sciobius2x2)
```

Format

A data frame with 47 observations on 25 grid cells.

Details

Original matrix is derived from Morrone (1994). An ulterior correction is introduced by Mast and Nyffeler (2003). We refer to this last version. Moreover, we adopt the same codification of species names.

Source

Morrone, J. J. 1994. *On the identification of areas of endemism.* Syst. Biol. 43:438-441.

Mast, A. R., and R. Nyffeler. 2003. *Using a null model to recognize significant co- occurrence prior to identifying candidate areas of endemism.* Syst. Biol. 52:271-280.

Examples

```
# Apply NAM analysis on Sciobius data set.
data(sciobius2x2)
fornam <- gridinfer(dntable = sciobius2x2)
namsc <- nam(fornam)
# Obtain the species belonging to each unit of co-occurrence
split(namsc$Categories[,1], namsc$Categories[,2])
```

---

track_OCP   *Clustering Behaviour along NAM*

Description

Evaluates the overall clustering performance of sub-networks along removal process.

Usage

```
track_OCP(histrval, msym)
```

Arguments

histrval	Integer vector. Index of last sub-network of node membership.
msym	A sympatry matrix.

Details

This function is auxilliary and is not intended to be called by the user.

Author(s)

Daniel A. Dos Santos

See Also

nam function makes a call to this function.

Index