Package ‘cml’

February 10, 2022

Type Package

Title  Conditional Manifold Learning

Version  0.0.2

Author  Anh Tuan Bui [aut, cre]

Maintainer  Anh Bui <atbui@u.northwestern.edu>

Depends  vegan, MASS

Description  Find a low-dimensional embedding of high-dimensional data, conditioning on auxiliary manifold information. The current version supports conditional MDS and conditional ISOMAP.

License  GPL-2

Encoding  UTF-8

RoxygenNote  6.0.1

NeedsCompilation  no

Repository  CRAN

Date/Publication  2022-02-10 09:30:02 UTC

R topics documented:

  cml-package .................................................. 2
  condDist .................................................... 3
  condIsomap .................................................. 3
  condMDS ..................................................... 5
  condSmacof ............................................... 6

Index  8
Description

Find a low-dimensional embedding of high-dimensional data, conditioning on auxiliary manifold information. The current version supports conditional MDS and conditional ISOMAP.


Details

Brief descriptions of the main functions of the package are provided below:

`condMDS()`: is the conditional MDS method, which uses conditional SMACOF to optimize its conditional stress function.

`condIsomap()`: is the conditional ISOMAP method, which is basically conditional MDS applying to graph distances (i.e., estimated geodesic distances) of the given distances/dissimilarities.

Author(s)

Anh Tuan Bui
Maintainer: Anh Tuan Bui <atbui@u.northwestern.edu>

References


Examples

```r
# generate toy data
N = 100
u.dim = 1
v.dim = 2
u.true = matrix(runif(N*u.dim), N, u.dim)
v.true = matrix(runif(N*v.dim), N, v.dim)
noise = rnorm(N, 0, .1)
d = dist(cbind(u.true, v.true, noise))

# Conditional MDS
u.cmds = condMDS(d, v.true, 1)$U

# Conditional ISOMAP
u.cisomap = condIsomap(d, v.true, 1, k = 20)$U
```
condDist

**conditional Euclidean distance**

**Description**

Internal functions for condSmacof_matrix() and condSmacof_vector().

**Usage**

```r
condDist(U, V.tilda, one_n_t=t(rep(1,nrow(U))))
condDist2(U, V.tilda2, one_n_t=t(rep(1,nrow(U))))
```

**Arguments**

- `U`: the embedding
- `V.tilda = V %*% B`
- `V.tilda2 = V %*% b^2*t(V)`
- `one_n_t = t(rep(1,nrow(U)))`

**Value**

A `dist` object.

**Author(s)**

Anh Tuan Bui

**References**


condIsomap

**conditional ISOMAP**

**Description**

Finds a low-dimensional manifold embedding of a given distance/dissimilarity matrix, conditioning on auxiliary manifold parameters. The method applies conditional MDS (see `condMDS`) to a graph distance matrix computed for the given distances/dissimilarities, using the `isomap` function.
condIsomap

Usage

condIsomap(d, V, u.dim, epsilon = NULL, k, W = NULL, 
method = c('matrix', 'vector'), 
it.max = 1000, gamma = 1e-05, 
init = c('none', 'user'), U.start = NULL, 
B.start = NULL, b.start = NULL, ...)

Arguments

d a distance/dissimilarity matrix of \(N\) entities (or a dist object).
V an \(N \times q\) matrix of \(q\) manifold auxiliary parameter values of the \(N\) entities.
u.dim the embedding dimension.
epsilon shortest dissimilarity retained.
k Number of shortest dissimilarities retained for a point. If both epsilon and k 
are given, epsilon will be used.
W an \(N \times N\) weight matrix. If not given, a matrix of ones will be used.
method if matrix, conditional SMACOF with a full \(B\) matrix will be used. If vector, 
conditional SMACOF with a diagonal \(B\) matrix will be used. The latter is more 
efficient for large \(q\).
it.max the max number of conditional SMACOF iterations.
gamma conditional SMACOF stops early if the reduction of normalized conditional 
stress is less than gamma
init method to initialize the starting values for the embedding and \(B\) (or \(b\)).
U.start user-defined starting values for the embedding (when \(U\).init = 'user')
B.start starting \(B\) matrix.
b.start starting diagonal values of the \(B\) matrix (which is restricted to be diagonal).
... other arguments for the isomap(vegan) function.

Value

U the embedding result.
B the estimated \(B\) matrix.
stress Normalized conditional stress value.
sigma the conditional stress value at each iteration.
init method to initialize the starting values for the embedding and \(B\) (or \(b\)).
U.start the starting values for the embedding.
B.start starting values for the \(B\) matrix, if method="matrix".
b.start the starting values of the \(B\) matrix (which is restricted to be diagonal), if method="vector".

Author(s)

Anh Tuan Bui
CondMDS

References

See Also
condMDS

Examples
# see help(cml)

condMDS

Conditional Multidimensional Scaling

Description
Finds a low-dimensional embedding of a given distance/dissimilarity matrix, conditioning on auxiliary manifold parameters. The method uses conditional SMACOF to optimize its conditional stress function.

Usage
condMDS(d, V, u.dim, W = NULL, method = c('matrix', 'vector'), it.max = 1000, gamma = 1e-05, init = c('none', 'user'), U.start = NULL, B.start = NULL, b.start = NULL)

Arguments
d a distance/dissimilarity matrix of N entities (or a dist object).
V an N x q matrix of q manifold auxiliary parameter values of the N entities.
u.dim the embedding dimension.
W an N x N weight matrix. If not given, a matrix of ones will be used.
method if matrix, conditional SMACOF with a full B matrix will be used. If vector, conditional SMACOF with a diagonal B matrix will be used. The latter is more efficient for large q.
it.max the max number of conditional SMACOF iterations.
gamma conditional SMACOF stops early if the reduction of normalized conditional stress is less than gamma
init method to initialize the starting values for the embedding and B (or b).
U.start user-defined starting values for the embedding (when U.init = 'user')
B.start starting B matrix.
b.start starting diagonal values of the B matrix (which is restricted to be diagonal).
Value

- \( U \) the embedding result.
- \( B \) the estimated \( B \) matrix.
- stress Normalized conditional stress value.
- sigma the conditional stress value at each iteration.
- init method to initialize the starting values for the embedding and \( B \) (or \( b \)).
- \( U.\text{start} \) the starting values for the embedding.
- \( B.\text{start} \) starting values for the \( B \) matrix, if method=\"matrix\".
- \( b.\text{start} \) the starting values of the \( B \) matrix (which is restricted to be diagonal), if method=\"vector\".

Author(s)

Anh Tuan Bui

References


See Also

condIsomap

Examples

```r
# see help(cml)
```

---

**condSmacof**

*Conditional SMACOF*

Description

Conditional SMACOF algorithms. Intended for internal usage.

Usage

```r
condSmacof_matrix(d, V, u.dim, W = NULL, it.max = 1000, gamma = 1e-05, init = c('none', 'user'), U.start, B.start)
condSmacof_vector(d, V, u.dim, W = NULL, it.max = 1000, gamma = 1e-05, init = c('none', 'user'), U.start, b.start)
```
condSmacof

Arguments

- **d**: a dist object of N entities.
- **V**: an Nxq matrix of q manifold auxiliary parameter values of the N entities.
- **u.dim**: the embedding dimension.
- **w**: an NxN weight matrix. If not given, a matrix of ones will be used.
- **it.max**: the max number of conditional SMACOF iterations.
- **gamma**: conditional SMACOF stops early if the reduction of normalized conditional stress is less than gamma.
- **init**: method to initialize the starting values for the embedding and B (or b).
- **U.start**: user-defined starting values for the embedding (when U.init = 'user')
- **B.start**: condSmacof_matrix() only; starting B matrix.
- **b.start**: for condSmacof_vector() only; starting diagonal values of the B matrix (which is restricted to be diagonal).

Details

condSmacof_matrix() uses a full B matrix, whereas condSmacof_vector() restricts B to be a diagonal matrix and uses vector operations on the diagonal elements of B. The latter is therefore faster when q is large.

Value

- **U**: the embedding result.
- **B**: the estimated B matrix.
- **stress**: Normalized conditional stress value.
- **sigma**: the conditional stress value at each iteration.
- **init**: method to initialize the starting values for the embedding and B (or b).
- **U.start**: the starting values for the embedding.
- **B.start**: the starting B matrix, if using condSmacof_matrix()
- **b.start**: the starting diagonal values of the B matrix (which is restricted to be diagonal), if using condSmacof_vector().

Author(s)

Anh Tuan Bui

References

Index

cml (cml-package), 2
  cml-package, 2
  condDist, 3
  condDist2 (condDist), 3
  condIsomap, 3, 6
  condMDS, 3, 5, 5
  condSmacof, 6
  condSmacof_matrix (condSmacof), 6
  condSmacof_vector (condSmacof), 6