Package ‘spatialTailDep’

February 20, 2015

Title Estimation of spatial tail dependence models

Description Provides functions implementing the pairwise M-estimator for parametric models for stable tail dependence functions described in ``An M-estimator of spatial tail dependence''
(Einmahl, J.H.J., Kiriliouk, A., Krajina, A. and Segers, J., 2014).
See http://arxiv.org/abs/1403.1975.

Version 1.0.2

Author Anna Kiriliouk, Johan Segers

Maintainer Anna Kiriliouk <anna.kiriliouk@uclouvain.be>

Depends R (>= 3.0.1), cubature, mvtnorm, SpatialExtremes

LazyData true

Collate 'tailIntEmp.r' 'gridUtilities.r' 'EstimationSmith.r'
'EstimationBR.r' 'Estimation.R' 'spatialTailDep-package.r'

NeedsCompilation yes

Repository CRAN

Date/Publication 2014-11-18 16:23:30

R topics documented:

AsymVar *Function* AsymVar

Description

Function to compute the asymptotic variance matrix of the pairwise M-estimator for the Smith model or the Brown-Resnick process.

Usage

```
AsymVar(pairs, model, theta, Tol = 1e-05)
```

Arguments

pairs	A q x 4 matrix giving the Cartesian coordinates of the q pairs of locations.
model	Choose between "smith" and "BR".
theta	Parameter vector. For the Smith model, theta must be equal to the 2 x 2 covariance matrix. For the Brown-Resnick pocess, theta $= (\alpha, \rho, \beta, c)$.
Tol	The tolerance in the numerical integration procedure. Defaults to 1e-5.

Details

For a matrix of coordinates of pairs of locations, this function returns the asymptotic variance matrix of the estimator. An optimal weight matrix can be defined as the inverse of the asymptotic variance matrix. For a detailed description of this procedure, see Einmahl et al. (2014).

The parameter vector theta must be a positive semi-definite matrix if model = "smith" and a vector of length four if model = "BR", where $0 < \alpha < 1$, $\rho > 0$, $0 < \beta \leq \pi/2$ and $c > 0$.

Value

A q x q matrix.

References

Einmahl, J.H.J., Kiriliouk, A., Krajina, A. and Segers, J. (2014), "An M-estimator of spatial tail dependence". See http://arxiv.org/abs/1403.1975.

See Also

Mestimator, selectPairIndices, pairCoordinates

Examples

```
## Define the locations of three stations
(locations <- rbind(c(1.0,1.0),c(2.0,1.0),c(1.2,2.5)))
## select pairs
(pairIndices <- selectPairIndices(locations, maxDistance = 3))
(pairs <- pairCoordinates(locations, pairIndices))
## Smith model parameter matrix
(theta <- rbind(c(1.5, .5), c(.5, 1)))
## The matrix. Takes a couple of seconds to compute.
## AsymVar(pairs, model = "smith", theta = theta, Tol = 1e-04)

## Parameters of the Brown-Resnick process
(theta <- c(1.5,1,0.5,0.25))
## The matrix. Takes a couple of seconds to compute.
## AsymVar(pairs, model = "BR", theta = theta, Tol = 1e-04)
```

KNMIdata *Wind speeds in the Netherlands.*

Description

Daily maximal speeds of wind gusts, measured in 0.1 m/s. The data are observed at 22 inland weather stations in the Netherlands. Only the summer months are presented here (June, July, August). Also included are the Euclidian coordinates of the 22 weather stations, where a distance of 1 corresponds to 100 kilometers. For more information on this dataset, see Einmahl et al. (2014).

Format

KNMIdata$data is a matrix with 672 rows and 22 columns, KNMI$loc is a matrix with 22 rows and 2 columns.

Source

http://www.knmi.nl/climatology/daily_data/selection.cgi

References

Einmahl, J.H.J., Kiriliouk, A., Krajina, A. and Segers, J. (2014), "An M-estimator of spatial tail dependence". See http://arxiv.org/abs/1403.1975.

Examples

```
data(KNMIdata)
locations <- KNMIdata$loc
pairIndices <- selectPairIndices(locations, maxDistance = 0.5)
Mestimator(KNMIdata$data, locations, pairIndices, k = 60, model="BR",
iterate=FALSE, covMat = FALSE)$theta
```

`Mestimator` *Function* `Mestimator`

Description

Function to compute the pairwise M-estimator for the parameters of the Smith model or the Brown-Resnick process.

Usage

```
Mestimator(x, locations, pairIndices, k, model, Tol = 1e-05,
  startingValue = NULL, Omega = diag(nrow(pairIndices)), iterate = TRUE,
  covMat = TRUE)
```

Arguments

`x`	An n x d data matrix.
`locations`	A d x 2 matrix containing the Cartesian coordinates of d points in the plane.
`pairIndices`	A q x 2 matrix containing the indices of q pairs of points from the matrix `locations`.
`k`	The threshold parameter in the definition of the empirical stable tail dependence function.
`model`	Choose between "smith" and "BR".
`Tol`	The tolerance parameter in the numerical integration procedure; defaults to 1e-05.
`startingValue`	Initial value of the parameters in the minimization routine. Defaults to diag(2) for the Smith model and (1, 1.5, 0.75, 0.75) for the BR process.
`Omega`	A q x q matrix specifying the metric with which the distance between the parametric and nonparametric estimates will be computed. The default is the identity matrix, i.e., the Euclidean metric.
`iterate`	A Boolean variable. If `TRUE` (the default), then the estimator is calculated twice, first with `Omega` specified by the user, and then a second time with the optimal `Omega` calculated at the initial estimate.
`covMat`	A Boolean variable. If `TRUE` (the default), the covariance matrix is calculated.

Details

For a detailed description of the estimation procedure, see Einmahl et al. (2014). Some tips for using this function:

- `n` versus `d`: if the number of locations d is small ($d < 8$ say), a sufficiently large sample size (eg $n > 2000$) is needed to obtain a satisfying result, especially for the Brown-Resnick process. However, if d is large, a sample size of $n = 500$ should suffice.

- `pairIndices`: if the number of pairs q is large, `Mestimator` will be rather slow. This is due to the calculation of `Omega` and `covMat`. Setting `iterate = FALSE` and `covMat = FALSE` will make this procedure fast even for several hundreds of pairs of locations.
- `Tol`: the tolerance parameter is used when calculating the three- and four-dimensional integrals in the asymptotic covariance matrix (see Appendix B in Einmahl et al. (2014)). A tolerance of 1e-04 often suffices, although the default tolerance is a safer choice.
- `StartingValue`: for the Smith model, the estimator usually doesn't depend on the starting value at all. For the Brown-Resnick process, it is advised to try a couple of starting values if d is very small, preferably a starting value with $c < 1$ and one with $c > 1$.
- `iterate`: if `iterate = TRUE`, the matrix `Omega` is calculated. This weight matrix tends to have a larger effect when d is large and/or when the Smith model is used.
- `covMat`: if the resulting covariance matrix is incorrect (eg negative diagonal values), then `Tol` is set too high. For the Smith model, the order of the parameters is $(\sigma_{11}, \sigma_{22}, \sigma_{12})$.

Value

A list with the following components:

`theta`	The estimator with estimated optimal weight matrix.
`theta_pilot`	The estimator without the optimal weight matrix.
`covMatrix`	The estimated covariance matrix for the estimator.
`Omega`	The weight matrix with which the estimator was calculated.

References

Einmahl, J.H.J., Kiriliouk, A., Krajina, A. and Segers, J. (2014), "An M-estimator of spatial tail dependence". See http://arxiv.org/abs/1403.1975.

See Also

`selectPairIndices`, `pairCoordinates`

Examples

```
## define the locations of 4 stations
(locations <- rbind(c(1,1),c(2,1),c(1,2),c(2,2)))
## select the pairs of locations; here, we select all locations
(pairIndices <- selectPairIndices(locations, maxDistance = 2))

## We use the rmaxstab function from the package SpatialExtremes to
## simulate from the Smith and the Brown-Resnick process.

## The Smith model
set.seed(2)
x<-rmaxstab(n = 5000, coord = locations,cov.mod="gauss",cov11=1,cov22=2,cov12=0.5)
## calculate the pairwise M-estimator. This may take up to one minute or longer.
## Mestimator(x, locations, pairIndices, 100, model="smith",Tol = 5e-04)

## The Brown-Resnick process
```

```
set.seed(2)
x <- rmaxstab(n = 5000, coord = locations, cov.mod = "brown", range = 3, smooth = 1)
## We can only simulate isotropic processes with rmaxstab, so we multiply the coordinates
## of the locations with V^(-1) (beta,c). Here we choose beta = 0.25 and c = 1.5
(Vmat<-matrix(c(cos(0.25),1.5*sin(0.25),-sin(0.25),1.5*cos(0.25)),nrow=2))
(locationsAniso <- locations %*% t(solve(Vmat)))
## calculate the pairwise M-estimator. This may take up to one minute or longer.
## Mestimator(x, locationsAniso, pairIndices, 300, model="BR",Tol = 5e-04)
```

`pairCoordinates` *Function* `pairCoordinates`

Description

Given a matrix of coordinates of locations and a matrix of indices of pairs of locations, returns a matrix with the coordinates of the pairs of locations.

Usage

```
pairCoordinates(locations, pairIndices)
```

Arguments

`locations`	A d x 2 matrix containing the Cartesian coordinates of d points in the plane.
`pairIndices`	A q x 2 matrix containing the indices of q pairs of points.

Value

A q x 4 matrix, where each row gives the Cartesian coordinates of the two locations in the corresponding pair.

See Also

`selectPairIndices`

Examples

```
(locations<-cbind(rep(1:2,3),rep(1:3,each=2)))
(pairs <- selectPairIndices(locations, maxDistance = 1.5))
pairCoordinates(locations, pairs)
```

selectPairIndices *Function* selectPairIndices

Description

Out of a list of locations given by their Cartesian coordinates, selects pairs of locations with a distance as a criterion.

Usage

```
selectPairIndices(locations, maxDistance = NULL, numberOfPairs = NULL)
```

Arguments

`locations`	A d x 2 matrix containing the Cartesian coordinates of d points in the plane.
`maxDistance`	Pairs of locations with distance not larger than `maxDistance` will be selected. If `NULL`, then the selection is made based on `numberOfPairs`.
`numberOfPairs`	The number of pairs that will be selected, taking distance as a criterion, with closer pairs being selected. It will return more pairs if there are several equidistant locations.

Value

A matrix with `q` rows and 2 columns, where `q` denotes the number of pairs selected. Each row contains the indices of the corresponding pair of selected locations in the input argument `locations`.

See Also

`pairCoordinates`

Examples

```
(locations<-cbind(rep(1:4,5),rep(1:5,each=4)))
selectPairIndices(locations, maxDistance = 1.5)
```

spatialTailDep *spatialTailDep*

Description

The package `spatialTailDep` provides functions implementing the pairwise M-estimator of parametric spatial tail dependence models for distributions attracted to a max-stable law. This is a rank-based estimator, constructed as the minimizer of the distance between a vector of integrals of parametric pairwise tail dependence functions and the vector of their empirical counterparts. It is especially suited for high-dimensional data since it relies on bivariate margins only and, as a consequence of the rank-based approach, the univariate marginal distributions need not be estimated. For a complete description of the pairwise M-estimator, see Einmahl et al. (2014).

Details

Currently, this package allows for estimation of the Brown-Resnick process and the Gaussian extreme value process (usually known as the Smith model) in two-dimensional space. The main function of this package is Mestimator, but several other functions are exported as well: tailInt returns the integral of a parametric bivariate stable tail dependence function over the unit square, tailIntEmp returns the integral of the bivariate empirical stable tail dependence function over the unit square, and AsymVar returns the asymptotic variance matrix for a list of pairs of locations defined by the user.

The function Mestimator combines these functions: first it computes a pilot estimator based on the Euclidian distance between the integrals of the parametric and the empirical stable tail dependence functions, then it calculates a weight matrix, defined as the inverse of the asymptotic variance matrix in the pilot estimator, and finally it returns the estimator obtained by replacing the Euclidian distance by a quadratic form based on the weight matrix. More details about this procedure can be found in Einmahl et al (2014).

The package exports two auxiliary functions as well: the function selectPairIndices returns a list of pair indices of locactions, based on either a maximum-distance criterion or on the maximal number of pairs that the user wants to include. Next, the function pairCoordinates can be used to select the pairs with these indices from a list of location coordinates.

References

Einmahl, J.H.J., Kiriliouk, A., Krajina, A. and Segers, J., "An M-estimator of spatial tail dependence". See http://arxiv.org/abs/1403.1975.

Examples

```
## get a list of all help files of user-visible functions in the package
help(package = spatialTailDep)
```

---

tailInt *Function* tailInt

Description

Integral of the bivariate parametric stable tail dependence function over the unit square, for the Smith model or the Brown-Resnick process.

Usage

```
tailInt(loc, model, theta)
```

Arguments

loc	A 2 x 2 matrix, where a row represents a location.
model	Choose between "smith" and "BR".
theta	Parameter vector. For the Smith model, theta must be equal to the 2 x 2 covariance matrix. For the Brown-Resnick pocess, theta $= (\alpha, \rho, \beta, c)$.

Details

This is an analytic implementation of the integral of the stable tail dependence function, which is much faster than numerical integration. For the definitions of the parametric stable tail dependence functions, see Einmahl et al. (2014).

The parameter vector `theta` must be a positive semi-definite matrix if `model = "smith"` and a vector of length four if `model = "BR"`, where $0 < \alpha < 1$, $\rho > 0$, $0 < \beta \leq \pi/2$ and $c > 0$.

Value

A scalar.

References

Einmahl, J.H.J., Kiriliouk, A., Krajina, A. and Segers, J. (2014), "An M-estimator of spatial tail dependence". See http://arxiv.org/abs/1403.1975.

See Also

`Mestimator`, `tailIntEmp`

Examples

```
tailInt(loc = cbind(c(1,1),c(2,3)), model = "smith", theta = rbind(c(3,1),c(1,2)))
tailInt(loc = cbind(c(1,2),c(3,4)), model = "BR", theta = c(1.5,1,0.5,0.75))
```

`tailIntEmp` *Function* `tailIntEmp`

Description

Integral of the bivariate empirical stable tail dependence function over the unit square.

Usage

```
tailIntEmp(ranks, n = nrow(ranks), k)
```

Arguments

`ranks`	A n x 2 matrix, where each column is a permutation of the integers `1:n`, representing the ranks computed from a sample of size `n`.
`n`	The sample size. If not specified, it is computed as the number of rows of the matrix `ranks`.
`k`	The threshold parameter in the definition of the empirical stable tail dependence function. An integer between 1 and `n-1`.

Details

This is an analytic implementation of the integral of the stable tail dependence function, which is much faster than numerical integration. See Einmahl et al. (2014) for a definition of the empirical stable tail dependence function.

Value

A scalar.

References

Einmahl, J.H.J., Kiriliouk, A., Krajina, A. and Segers, J. (2014), "An M-estimator of spatial tail dependence". See http://arxiv.org/abs/1403.1975.

See Also

Mestimator, tailInt

Examples

```
n <- 20
(ranks <- cbind(sample(1:n), sample(1:n)))
tailIntEmp(ranks, k = 5)
```

Index