

Package ‘SpatialNP’

October 16, 2009

Type Package

Title Multivariate nonparametric methods based on spatial signs and ranks

Version 1.0-1

Date 2009-10-16

Author Seija Sirkia, Klaus Nordhausen, Hannu Oja

Maintainer Seija Sirkia <seija.sirkia@iki.fi>

Depends ICSNP

Description This package contains test and estimates of location, tests of independence, tests of sphericity and several estimates of shape all based on spatial signs, symmetrized signs, ranks and signed ranks.

License GPL-2

Repository CRAN

Date/Publication 2009-10-16 12:07:22

R topics documented:

SpatialNP-package	2
Independence tests	2
Location tests	4
Shape matrices	6
Spatial location	7
Spatial sign and rank covariance matrices	9
Spatial symmetrized signs, ranks and signed ranks	10
Sphericity tests	11
to.shape	13
Index	15

SpatialNP-package *Multivariate nonparametric tests and estimates based on spatial signs and ranks.*

Description

This package contains tests and an estimate of location, tests of independence, tests of sphericity and several estimates of shape all based on spatial signs, symmetrized signs (that is, signs of pairwise differences), ranks and signed ranks.

Details

Package: SpatialNP
Type: Package
Version: 1.0
Date: 2008-07-02
License: GPL (>= 2)

In this package there are three functions for inference, [sr.loc.test](#), [sr.indep.test](#) and [sr.sphere.test](#), for location, independence and sphericity tests. The so called **inner** and **outer** standardization matrices are also available as well as the actual **sign and rank** score functions, together with a utility function [to.shape](#).

Author(s)

Seija Sirkia, Klaus Nordhausen, Hannu Oja Maintainer: Seija Sirkia, seija.sirkia@iki.fi

References

Oja, H., Randles R. (2004) *Multivariate Nonparametric Tests*. Statistical Science 19, 598-605

See Also

package [ICSNP](#)

Independence tests *Multivariate test of independence based on spatial signs or ranks*

Description

Test of independence between two sets of variables. Inference is based on the spatial signs of the observations, symmetrized signs of the observations or spatial signed ranks of the observations.

Usage

```
sr.indep.test(X, Y = NULL, g = NULL, score = c("sign",
"symmsign", "rank"), regexp= FALSE, cond = FALSE, cond.n = 1000,
na.action = na.fail)
```

Arguments

<code>X</code>	a matrix or a data frame
<code>Y</code>	an optional matrix or a data frame
<code>g</code>	a factor giving the two sets of variables, or numeric vector or vector of column names giving the first set of variables. See details
<code>score</code>	a character string indicating which transformation of the observations should be used
<code>regexp</code>	logical. Is <code>g</code> a regular expression?
<code>cond</code>	logical. Should the conditionally distribution free test be used?
<code>cond.n</code>	Number of permutations to use in the conditionally distribution free test
<code>na.action</code>	a function which indicates what should happen when the data contain 'NA's. Default is to fail.

Details

`X` should contain the first set of variables and `Y` the second with matching rows. Alternatively, `X` should contain both sets and `g` should be a factor of length equal to number of columns of `X`, or, `g` should be a numeric or character vector naming the variables in the first set. If `g` is a character vector it is assumed to name all wanted columns exactly, unless `regexp` is `TRUE`.

Value

A list with class 'htest' containing the following components:

<code>statistic</code>	the value of the statistic
<code>parameter</code>	the degrees of freedom for the statistic or the number of replications if conditionally distribution free p-value was used
<code>p.value</code>	the p-value for the test
<code>null.value</code>	the specified hypothesized value of the measure of dependence (always 0)
<code>alternative</code>	a character string with the value 'two.sided'.
<code>method</code>	a character string indicating what type of test was performed
<code>data.name</code>	a character string giving the name of the data (and grouping vector)

Author(s)

Seija Sirkia, <seija.sirkia@iki.fi>

See Also

[Spatial signs and ranks](#)

Examples

```

A<-matrix(c(1,2,-3,4,3,-2,-1,0,4),ncol=3)
X<-matrix(rt(150,1),ncol=3)%*%t(A)
Y<-cbind(X+runif(150,-1,1),runif(50))
sr.indep.test(X,Y)
#alternative calls:
Z<-cbind(X,Y)
colnames(Z)<-c("a1","a2","a3","b1","b2","b3","b4")
g<-factor(c(rep(1,3),rep(2,4)))
sr.indep.test(Z,g=g)
sr.indep.test(Z,g=c("b"),regex=TRUE)
sr.indep.test(Z,g=1:3)

```

Location tests

Spatial sign and rank tests of multivariate location

Description

Multivariate tests of location of one or more samples based on spatial signs and (signed) ranks. In case of one sample the null hypothesis about a given location is tested. In case of several samples the null hypothesis is that all samples have the same location.

Usage

```

sr.loc.test(X, Y = NULL, g = NULL, score = c("sign", "rank"),
nullvalue = NULL, cond = FALSE, cond.n = 1000, na.action = na.fail,
...)

```

Arguments

X	a matrix or a data frame
Y	an optional matrix or a data frame
g	a factor giving the groups (may contain just one level)
score	a character string indicating which transformation of the observations should be used
nullvalue	location to be tested in the one sample case (ignored if there is more than one sample)
cond	logical. Should the conditionally distribution free test be used? (Ignored if score is "rank")
cond.n	number of permutations to use in the conditionally distribution free test
na.action	a function which indicates what should happen when the data contain 'NA's. Default is to fail.
...	further arguments to be passed to other functions

Details

X should contain the the whole data set and g should describe the groups, or, if there is only one group, g may be missing. Alternatively, if there are two samples X may contain only the first sample while the second sample is given in Y and g is ignored. Note that in the one sample case when rank is chosen as score the function in fact uses signed ranks.

Note that the conditionally distribution free p-value is only provided for the sign based version of the test.

Value

A list with class 'htest' containing the following components:

statistic	the value of the statistic
parameter	the degrees of freedom for the statistic or the number of replications if conditionally distribution free p-value was used
p.value	the p-value for the test
null.value	the specified hypothesized value of the (common) location
alternative	a character string with the value 'two.sided'.
method	a character string indicating what type of test was performed
data.name	a character string giving the name of the data (and grouping vector)

Author(s)

Seija Sirkia, (seija.sirkia@iki.fi)

See Also

[Spatial signs and ranks](#), [HotellingsT2](#) for the classical Hotelling's T^2 test

Examples

```
A<-matrix(c(1,2,-3,4,3,-2,-1,0,4),ncol=3)
X<-rbind(matrix(rt(100*3,1),ncol=3),matrix(rt(50*3,1)+1,ncol=3))%*%t(A)
sr.loc.test(X,cond=TRUE)
X[1:50,]<-X[1:50,]+1
g<-factor(rep(c(1,2,3),each=50))
sr.loc.test(X,g=g,score="rank")
```

Shape matrices *Shape matrices based on spatial ranks and signed ranks*

Description

Iterative algorithms to find shape matrices based on spatial signs and ranks and the k-step versions of these.

Usage

```
rank.shape(X, init = NULL, steps = Inf, eps = 1e-06, maxiter = 100,
na.action = na.fail)
```

```
signrank.shape(X, location = NULL, init = NULL, steps = Inf, eps =
1e-06, maxiter = 100, na.action = na.fail)
```

```
spatial.shape(X, score = c("sign", "symmsign", "rank", "signrank"),
location = NULL, init = NULL, steps = Inf, eps = 1e-06, maxiter = 100,
na.action = na.fail)
```

Arguments

<code>X</code>	a matrix or a data frame
<code>score</code>	a character string indicating which transformation of the observations should be used
<code>location</code>	an optional vector giving the location of the data
<code>init</code>	an optional starting value for the iteration
<code>steps</code>	fixed number of iteration steps to take, if <code>Inf</code> iteration is repeated until convergence (or until <code>maxiter</code> steps)
<code>eps</code>	tolerance for convergence
<code>maxiter</code>	maximum number of iteration steps. Ignored if <code>steps</code> is finite
<code>na.action</code>	a function which indicates what should happen when the data contain 'NA's. Default is to fail.

Details

As [tyler.shape](#) for spatial signs and [duembgen.shape](#) for spatial symmetrized signs, but for spatial ranks and signed ranks. These are the so called inner standardization matrices of location etc. tests based on spatial signs and ranks. When data is standardized using these matrices the corresponding sign or rank scores will appear “uncorrelated”: the corresponding outer standardization matrices will be proportional to the identity matrix, see examples.

`spatial.shape` is a wrapper function for a unified access to all four shape estimates. The choice of estimate is done via `score`:

- "sign" for `tyler.shape`
- "symmsign" for `duembgen.shape`
- "rank" for `rank.shape`
- "signrank" for `signrank.shape`

`signrank.shape` (and `tyler.shape`, thus also `spatial.shape`) requires the location vector with respect to which it is computed. If none is provided, vector of column means is used.

Author(s)

Seija Sirkia, <seija.sirkia@iki.fi>

See Also

[tyler.shape](#), [duembgen.shape](#), also [spatial sign and rank covariance matrices](#) and [spatial signs and ranks](#)

Examples

```
A<-matrix(c(1,2,-3,4,3,-2,-1,0,4),ncol=3)
X<-matrix(rt(150,1),ncol=3)%*%t(A)
signrank.shape(X)
spatial.shape(X,score="sign")
to.shape(A%*%t(A))
# one-step shape estimate based on spatial ranks and covariance matrix:
spatial.shape(X,score="rank",init=cov(X),steps=1)
# effect of inner standardization:
RCov(X)
RCov(X%*%t(chol(solve(rank.shape(X)))))
```

Spatial location *Multivariate location estimates based on spatial signs and signed ranks*

Description

Spatial median, multivariate Hodges-Lehmann estimate of location and their affine equivariant versions

Usage

```
ae.hl.estimate(X, init=NULL, shape=TRUE, maxiter = 500, eps = 1e-6,
na.action = na.fail)
```

```
spatial.location(X, score = c("sign", "signrank"), init = NULL, shape
= TRUE, maxiter = 500, eps = 1e-6, na.action = na.fail)
```

Arguments

<code>X</code>	a matrix or a data frame
<code>score</code>	a character string indicating which transformation of the observations should be used
<code>init</code>	an optional vector giving the initial point of the iteration
<code>shape</code>	logical, or a matrix. See details
<code>eps</code>	tolerance for convergence
<code>maxiter</code>	maximum number of iteration steps
<code>na.action</code>	a function which indicates what should happen when the data contain 'NA's. Default is to fail.

Details

Hodges-Lehmann estimate of multivariate location is the spatial median (see [spatial.median](#)) of the observations and their pairwise averages. Since spatial median is not affine equivariant, neither is multivariate Hodges-Lehmann estimate. Affine equivariance can be achieved by simultaneously estimating the corresponding shape, as proposed for the spatial median by Hettmansperger and Randles (2002). For spatial median the corresponding shape is [tyler.shape](#) and the simultaneous estimation is implemented as [HR.Mest](#), see its documentation for further details on the algorithm. For the Hodges-Lehmann estimate it is [rank.shape](#).

`spatial.location` is a wrapper function for a unified access to both location estimates. The choice of estimate is done via `score`:

- "sign" for spatial median
- "signrank" for Hodges-Lehmann estimate

If a matrix (must be symmetric and positive definite, but this is not checked) is given as `shape` the location estimate is found with respect to that shape and no further shape estimation is done. If a logical TRUE is given as `shape` the shape is estimated and consequently the affine equivariant version of the location estimate is found. If `shape` is FALSE then shape estimation is not done and the non affine equivariant version of the location estimate is found.

Value

The estimate vector with the (final estimate of or given) shape matrix as attribute "shape".

Author(s)

Seija Sirkia, <seija.sirkia@iki.fi>

References

Hettmansperger, T. and Randles, R. (2002) A Practical Affine Equivariant Multivariate Median, *Biometrika*, 89, pp. 851-860

See Also

[spatial.median](#), [HR.Mest](#) [signrank.shape](#)

Examples

```
A<-matrix(c(1,2,-3,4,3,-2,-1,0,4),ncol=3)
X<-matrix(rt(150,1),ncol=3)%*%t(A)
spatial.location(X,score="signrank")
spatial.location(X,score="sign")
#compare with:
colMeans(X)
ae.hl.estimate(X,shape=A%*%t(A))
ae.hl.estimate(X,shape=FALSE)
```

Spatial sign and rank covariance matrices

Spatial sign and rank covariance matrices

Description

Functions to compute spatial sign, spatial symmetrized sign, spatial rank and spatial signed rank covariance matrices

Usage

```
SCov(X, location, na.action = na.fail)
SSCov(X, na.action = na.fail)
RCov(X, na.action = na.fail)
SRCov(X, location, na.action = na.fail)
```

Arguments

X	matrix or a data frame
location	numeric vector (may be missing)
na.action	a function which indicates what should happen when the data contain 'NA's. Default is to fail.

Details

These functions compute the matrices of the form

$$ave\{S(x_i)S^T(x_i)\}$$

where $S(x_i)$ are the appropriate scores of the data: spatial signs, spatial symmetrized signs, spatial ranks or spatial signed ranks. These are the so called outer standardization matrices of location etc. tests based on spatial signs and ranks. They are not affine equivariant.

SCOV and SRCOV require a location vector with respect to which they are computed. If none is provided, vector of column means is used.

Author(s)

Seija Sirkia, <seija.sirkia@iki.fi>

See Also

[spatial signs and ranks, corresponding shape matrices](#) (inner standardization matrices)

Examples

```
A<-matrix(c(1,2,-3,4,3,-2,-1,0,4),ncol=3)
X<-matrix(rt(150,1),ncol=3)%*%t(A)
SCov(X)
SSCov(X)
RCov(X)
SRCov(X)
to.shape(A%*%t(A),trace=1)
```

Spatial symmetrized signs, ranks and signed ranks
Spatial symmetrized signs, ranks and signed ranks

Description

Functions to compute spatial symmetrized signs, ranks and signed ranks.

Usage

```
spatial.symmsign(X, shape = TRUE, na.action = na.fail, ...)
spatial.rank(X, shape = TRUE, na.action = na.fail, ...)
spatial.signrank(X, center = TRUE, shape = TRUE, na.action = na.fail, ...)
```

Arguments

X	a matrix or a data frame
center	a vector or a logical, see details
shape	a matrix or a logical, see details
na.action	a function which indicates what should happen when the data contain 'NA's. Default is to fail.
...	arguments that can be passed on to function used for the estimation of shape.

Details

The spatial signs of an observed vector is simply the vector, possibly affinely transformed first, multiplied by its Euclidian length. See [spatial.sign](#) for a precise definition. Symmetrized spatial signs are the spatial signs of the pairwise differences of the data

$$\|x_i - x_j\|^{-1}(x_i - x_j)$$

(there are n over 2 of these). Spatial rank of an observation is the average of the signs of the differences of that observation and the others:

$$R(x_i) = \text{ave}_j\{\|x_i - x_j\|^{-1}(x_i - x_j)\}$$

Spatial signed rank of an observation is defined as

$$Q(x_i) = (R(x_i) + ave_j\{|x_i + x_j|^{-1}(x_i + x_j)\})/2$$

If a numerical value is given for `shape` and/or `center` these are used to transform the data before the computation of signs or ranks. A logical `TRUE` indicates that the shape or center should be estimated. In this case an affine transformation that makes the resulting signs or ranks have a covariance matrix equal or proportional to the identity matrix and centered on the origin is found. A logical `FALSE` indicates that the null value, that is, the identity matrix or the origin, should be used. Note that only signed ranks depend on a center.

The value of `shape` and/or location used are returned as attributes.

Author(s)

Seija Sirkia, <seija.sirkia@iki.fi>

See Also

[spatial.sign](#) for the signs, [spatial sign and rank matrices](#) and [ae.hl.estimate](#) for the standardizing transformations

Examples

```
A<-matrix(c(1,2,-3,4),ncol=2)
X<-matrix(rt(20,1),ncol=2)%*%t(A)
def.par<-par(no.readonly=TRUE) # for resetting
layout(matrix(1:4,ncol=2,nrow=2,byrow=TRUE))
plot(X,col=c(2,rep(1,19)))
plot(spatial.symmsign(X),col=c(2,rep(1,19)),xlim=c(-1,1),ylim=c(-1,1))
theta<-seq(0,2*pi,length=1000)
lines(sin(theta),cos(theta))
plot(spatial.rank(X),col=c(2,rep(1,19)),xlim=c(-1,1),ylim=c(-1,1))
lines(sin(theta),cos(theta))
plot(spatial.signrank(X),col=c(2,rep(1,19)),xlim=c(-1,1),ylim=c(-1,1))
lines(sin(theta),cos(theta))
par(def.par)
```

Sphericity tests *Sphericity tests*

Description

Tests of sphericity based on spatial signs and spatial signs of pairwise differences.

Usage

```
sr.sphere.test(X, score = c("sign", "symmsign"), shape = NULL,
na.action = na.fail)
```

Arguments

<code>X</code>	a matrix or a data frame
<code>score</code>	a character string indicating which transformation of the observations should be used
<code>shape</code>	a matrix with which the data should be standardized before the sphericity test
<code>na.action</code>	a function which indicates what should happen when the data contain 'NA's. Default is to fail.

Details

The test is for a null hypothesis of the form “true shape matrix is equal to the identity matrix”. Effectively, giving a matrix as `shape` will produce a test of whether the true shape is equal (in fact, proportional, since the scale of `shape` will have no effect) to it. In that case the test will still be for sphericity but the data is standardized beforehand.

Value

A list with class 'hstest' containing the following components:

<code>statistic</code>	the value of the statistic
<code>parameter</code>	the degrees of freedom for the statistic
<code>p.value</code>	the p-value for the test
<code>null.value</code>	the specified hypothesized value of the shape (always $\text{diag}(p)$ where p is the number of dimensions)
<code>alternative</code>	a character string with the value 'two.sided'.
<code>method</code>	a character string indicating what type of test was performed
<code>data.name</code>	a character string giving the name of the data

Author(s)

Seija Sirkia, <seija.sirkia@iki.fi>

See Also

[sign and rank covariance matrices](#)

Examples

```
A<-matrix(c(1,2,-3,4,3,-2,-1,0,4),ncol=3)
X<-matrix(rt(150,1),ncol=3)%*%t(A)
sr.sphere.test(X,score="sign")
```

to.shape *Rescale a matrix to a shape matrix*

Description

This function rescales a given matrix such that its determinant, trace or the value of the first diagonal element meets a criteria.

Usage

```
to.shape(M, determ, trace, first)
```

Arguments

M	a matrix to be scaled
determ	required value for determinant
trace	required value for trace
first	required value of the first diagonal element

Details

If `determ`, `trace` or `first` is given `M` is scaled such that its determinant, trace or first diagonal element, respectively, equals that value. If none of the three is given `M` is scaled such that its determinant equals one. If more than one criteria is given the first of them is used and the others silently ignored.

Value

The rescaled matrix

Note

A shape matrix is symmetric and positive definite square matrix. In order for the result to be such the argument matrix `M` should also be symmetric and positive definite square matrix. However, the function does not check for this. Expect to see errors if `M` is of inappropriate type.

Author(s)

Seija Sirkia, <seija.sirkia@iki.fi>

References

Paindaveine D. (2008) A Canonical Definition of Shape. *Statistics and Probability Letters* 78, 2240-2247

Examples

```
(A<-matrix(c(1,2,-3,4,3,-2,-1,0,4),ncol=3))  
to.shape(A**t(A))  
to.shape(A**t(A),trace=3)  
to.shape(A**t(A),first=1)
```

Index

*Topic **h**test

- Independence tests, 2
- Location tests, 4
- Shape matrices, 5
- Sphericity tests, 11

*Topic **m**ultivariate

- Independence tests, 2
- Location tests, 4
- Shape matrices, 5
- Spatial location, 7
- Spatial sign and rank
 - covariance matrices, 9
- Spatial symmetrized signs,
 - ranks and signed ranks, 10
- Sphericity tests, 11
- to.shape, 13

*Topic **n**onparametric

- Independence tests, 2
- Location tests, 4
- Shape matrices, 5
- Spatial location, 7
- Spatial sign and rank
 - covariance matrices, 9
- Spatial symmetrized signs,
 - ranks and signed ranks, 10
- Sphericity tests, 11

*Topic **p**ackage

- SpatialNP-package, 1

ae.hl.estimate, 11

ae.hl.estimate(*Spatial location*), 7

corresponding shape matrices, 9

duembgen.shape, 6

HotellingsT2, 5

HR.Mest, 8

ICSNP, 2

Independence tests, 2

inner, 2

Location tests, 4

outer, 2

rank.shape, 8

rank.shape(*Shape matrices*), 5

RCov(*Spatial sign and rank covariance matrices*), 9

SCov(*Spatial sign and rank covariance matrices*), 9

Shape matrices, 5

sign and rank, 2

sign and rank covariance matrices, 12

signrank.shape, 8

signrank.shape(*Shape matrices*), 5

Spatial location, 7

Spatial sign and rank covariance matrices, 9

spatial sign and rank covariance matrices, 6

spatial sign and rank matrices, 11

Spatial signs and ranks, 3, 5

spatial signs and ranks, 6, 9

Spatial symmetrized signs,

- ranks and signed ranks, 10

spatial.location(*Spatial location*), 7

spatial.median, 8

spatial.rank(*Spatial symmetrized signs, ranks and signed ranks*), 10

spatial.shape(*Shape matrices*), 5

spatial.sign, 10, 11

`spatial.signrank`(*Spatial
symmetrized signs, ranks
and signed ranks*), 10

`spatial.symmsign`(*Spatial
symmetrized signs, ranks
and signed ranks*), 10

`SpatialNP` (*SpatialNP-package*), 1

`SpatialNP-package`, 1

Sphericity tests, 11

`sr.indep.test`, 2

`sr.indep.test` (*Independence
tests*), 2

`sr.loc.test`, 2

`sr.loc.test` (*Location tests*), 4

`sr.sphere.test`, 2

`sr.sphere.test` (*Sphericity
tests*), 11

`SRCov` (*Spatial sign and rank
covariance matrices*), 9

`SSCov` (*Spatial sign and rank
covariance matrices*), 9

`to.shape`, 2, 13

`tyler.shape`, 6, 8