Package 'sscor'

April 29, 2025

Type Package

Title Robust Correlation Estimation and Testing Based on Spatial Signs

Version 0.2.1

Date 2025-04-29

Author Alexander Duerre [aut, cre],
Daniel Vogel [aut]

Maintainer Alexander Duerre <alexander.duerre@tu-dortmund.de>

Depends pcaPP, robustbase, mvtnorm

Description Provides the spatial sign correlation and the two-
stage spatial sign correlation as well as a one-sample test for the correlation coefficient.

License GPL-2 | GPL-3

ByteCompile true

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2025-04-29 13:10:02 UTC

Contents

---

sscor-package   *Correlation estimation by spatial signs*

Description

R functions for correlation estimation based on spatial signs. Including spatial sign correlation and two stage spatial sign correlation.

Details

Package:	sscor
Type:	Package
Version:	0.1
Date:	2015-4-22
License:	GPL-2	GPL-3

Author(s)

Alexander Dürre, Daniel Vogel

Maintainer: Alexander Dürre <alexander.duerre@udo.edu>

References

Dürre, A., Vogel, D., Fried, R. (2015): Spatial sign correlation, *Journal of Multivariate Analyis*, vol. 135, 89–105. arvix 1403.7635

Dürre, A., Vogel, D. (2016): Asymptotics of the two-stage spatial sign correlation, *Journal of Multivariate Analyis*, vol. 144, 54–67. arxiv 1506.02578

Dürre, A., Tyler, D. E., Vogel, D. (2016): On the eigenvalues of the spatial sign covariance matrix in more than two dimensions, to appear in: *Statistics and Probability Letters*. arvix 1512.02863

---

evShape2evSSCM   *Calculation of the eigenvalues of the Spatial Sign Covariance Matrix*

Description

evShape2evSSCM transforms the eigenvalues of the shape matrix of an elliptical distribution into that of the spatial sign covariance matrix.

Usage

```
evShape2evSSCM(evShape)
```

Arguments

evShape   (required) p-dimensional numeric, representing the eigenvalues of the shape matrix.

Details

The eigenvalues of the SSCM can be calculated from the eigenvalues of the shape matrix by numerical evaluation of onedimensional integrals, see Proposition 3 of Dürre, Tyler, Vogel (2016). We use the substitution

$$x = \frac{1+t}{1-t}$$

and Gaussian quadrature with Jacobi polynomials up to order 500 and $\beta = 0$ as well as $\alpha = p/2 - 1$, see chapter 2.4 (iv) of Gautschi (1997) for details.

The nodes and weights of the Gauss-Jacobi-quadrature are originally computed by the `gaussquad` package and saved in the file `jacobiquad` for faster computation.

Value

p-dimensional numeric, representing the eigenvalues of the corresponding spatial sign covariance matrix.

References

Dürre, A., Vogel, D., Fried, R. (2015): Spatial sign correlation, *Journal of Multivariate Analyis*, vol. 135, 89–105. arvix 1403.7635

Dürre, A., Tyler, D. E., Vogel, D. (2016): On the eigenvalues of the spatial sign covariance matrix in more than two dimensions, to appear in: *Statistics and Probability Letters*. arvix 1512.02863

Gautschi, W. (1997): *Numerical Analysis - An Introduction*, Birkhäuser, Basel.

Novomestky, F. (2013): gaussquad: Collection of functions for Gaussian quadrature. R package version 1.0-2.

See Also

Calculating the theoretical SSCM from the theoretical shape matrix Shape2SSCM

Examples

```
# defining eigenvalues of the shape matrix
evShape <- seq(from=0,to=1,by=0.1)

# standardized to have sum 1
evShape <- evShape/sum(evShape)

# calculating the related eigenvalues of the SSCM
evSSCM <- evShape2evSSCM(evShape)

plot(evShape,evSSCM)

# recalculate the eigenvalues of the shape matrix
```

```
evShape2 <- evSSCM2evShape(evSSCM)

# error is negligible
sum(abs(evShape-evShape2))
```

---

evSSCM2evShape   *Calculation of the eigenvalues of the shape matrix*

Description

evSSCM2evShape transforms the eigenvalues of the SSCM of an elliptical distribution into that of
the shape matrix.

Usage

```
evSSCM2evShape(delta,tol=10^(-10),itermax=100)
```

Arguments

delta   (required) p-dimensional numeric representing the eigenvalues of the SSCM.

tol   (optional) numeric, defines the stopping rule of the approximation procedure,
                 see details.

itermax   (optional) numeric, defines the maximal number of iterations, see details.

Details

The eigenvalues of the SSCM given that of the shape matrix can be calculated by evaluations of
numerical integrals, see the help of evShape2evSSCM or Dürre, Tyler, Vogel (2016). There is no
expression for the inverse relationshop known. Though one can apply a fixed point iteration to get
an approximation of the eigenvalues of the shape matrix. The iteration stops if either the maximal
number of iterations is reached, which produces a warning, or if the error between the eigenvalues
of the SSCM and the ones calculated from the actual fixed point iteration in L1 norm is smaller
than the given tolerance. Since the mapping between the sets of eigenvalues is injective, see Dürre,
Tyler, Vogel (2016), this gives a reasonable approximation of the eigenvalues of the shape matrix.

Value

p-dimensional numerical, representing the eigenvalues of the shape matrix. They are standardized
to sum to 1.

References

Dürre, A., Vogel, D., Fried, R. (2015): Spatial sign correlation, *Journal of Multivariate Analyis*,
vol. 135, 89–105. arvix 1403.7635

Dürre, A., Tyler, D. E., Vogel, D. (2016): On the eigenvalues of the spatial sign covariance matrix
in more than two dimensions, to appear in: *Statistics and Probability Letters*. arvix 1512.02863

See Also

Calculating the theoretical shape from the theoretical SSCM SSCM2Shape

Calculating the eigenvalues of the SSCM from the eigenvalues of the shape matrix evShape2evSSCM

Examples

```
# defining eigenvalues of the shape matrix
evShape <- seq(from=0,to=1,by=0.1)

# standardized to have sum 1
evShape <- evShape/sum(evShape)

# calculating the related eigenvalues of the SSCM
evSSCM <- evShape2evSSCM(evShape)

plot(evShape,evSSCM)

# recalculate the eigenvalues of the shape matrix
evShape2 <- evSSCM2evShape(evSSCM)

# error is negligible
sum(abs(evShape-evShape2))
```

Shape2SSCM   *Calculation of the Spatial Sign Covariance Matrix*

Description

Shape2SSCM transforms the theoretical shape matrix of an elliptical distribution into the spatial sign covariance matrix.

Usage

```
Shape2SSCM(V)
```

Arguments

V   (required) p x p matrix representing the theoretical shape matrix.

Details

The calculation consists of three steps. First one calculates eigenvectors and eigenvalues of the shape matrix by the function eigen. Then one determines the related eigenvalues of the SSCM using the function evShape2evSSCM and finally one expands the resulting eigendecomposition consisting of the eigenvectors of the Shape matrix and the eigenvalues of the SSCM. Note that this procedure only works for elliptical distributions.

Value

p x p symmetric numerical matrix, representing the spatial sign covariance matrix, which corresponds to the given shape matrix.

References

Dürre, A., Vogel, D., Fried, R. (2015): Spatial sign correlation, *Journal of Multivariate Analyis*, vol. 135, 89–105. arvix 1403.7635

Dürre, A., Tyler, D. E., Vogel, D. (2016): On the eigenvalues of the spatial sign covariance matrix in more than two dimensions, to appear in: *Statistics and Probability Letters*. arvix 1512.02863

See Also

Calculating the theoretical shape from the theoretical SSCM SSCM2Shape

Calculating the eigenvalues of the SSCM evShape2evSSCM

Examples

```
# defining a shape matrix with trace 1
V <- matrix(c(1,0.8,-0.2,0.8,1,0,-0.2,0,1),ncol=3)/3
V

# calculating the related SSCM
SSCM <- Shape2SSCM(V)

# recalculate the shape based on the SSCM
V2 <- SSCM2Shape(SSCM)
V2

# error is negligible
sum(abs(V-V2))
```

SSCM2Shape   *Calculation of the shape matrix*

Description

SSCM2Shape transforms the spatial sign covariance matrix of an elliptical distribution into its standardized shape matrix.

Usage

```
SSCM2Shape(V,itermax=100,tol=10^(-10))
```

Arguments

V	(required) p x p matrix representing the theoretical SSCM.
tol	(optional) numeric, defines the stopping rule of the approximation procedure, see the help of evSSCM2evShape for details.
itermax	(optional) numeric, defines the maximal number of iterations, see the help of evSSCM2evShape for details.

Details

The calculation consists of three steps. First one calculates eigenvectors and eigenvalues of the SSCM matrix by the function eigen. Then one determines the eigenvalues of the related Shape matrix using the function evSSCM2evShape. Finally one expands the eigendecomposition consisting of the eigenvectors of the SSCM and the eigenvalues of the shape matrix. The resulting shape matrix is standardized to have a trace of 1. Note that this procedure only works for elliptical distributions.

Value

p x p symmetric numerical matrix, representing the shape matrix with trace 1, which corresponds to the spatial sign covariance matrix.

References

Dürre, A., Vogel, D., Fried, R. (2015): Spatial sign correlation, *Journal of Multivariate Analyis*, vol. 135, 89–105. arvix 1403.7635

Dürre, A., Tyler, D. E., Vogel, D. (2016): On the eigenvalues of the spatial sign covariance matrix in more than two dimensions, to appear in: *Statistics and Probability Letters*. arvix 1512.02863

See Also

Calculating the theoretical shape from the theoretical SSCM SSCM2Shape

Calculating the eigenvalues of the SSCM evShape2evSSCM

Examples

```
# defining a shape matrix with trace 1
V <- matrix(c(1,0.8,-0.2,0.8,1,0,-0.2,0,1),ncol=3)/3
V

# calculating the related SSCM
SSCM <- Shape2SSCM(V)

# recalculate the shape based on the SSCM
V2 <- SSCM2Shape(SSCM)
V2

# error is negligible
sum(abs(V-V2))
```

---

sscor   *Spatial sign correlation*

Description

sscor computes a robust correlation matrix estimate based on spatial signs, as described in Dürre et al. (2015).

Usage

```
sscor(X, location=c("2dim-median","1dim-median","pdim-median","mean"),
scale=c("mad","Qn","sd"), standardized=TRUE, pdim=FALSE, ...)
```

Arguments

X	(required) p x n data matrix, number of colums is the dimension p and the number of rows is the number of observations n.
location	(optional) either a p-dimensional numeric vector specifying the location or a character string indicating the location estimator to be used. Possible values are "2dim-median","1dim-median","pdim-median","mean". The default is "2dim-median". See details below.
scale	(optional) either a p-dimensional numeric vector specifying the p marginal scales or a character string indicating the scale estimator to be used. Possible values are "mad","Qn","sd". The default is "mad". See details below.
standardized	(optional) logical; indicating whether the data should be standardized by marginal scale estimates prior to computing the spatial sign correlation. The default is TRUE.
pdim	(optional) logical; indicating whether the correlation matrix consists of pairwise correlation estimates or is estimated at once by the p-dimensional spatial sign correlation, see details.
...	(optional) arguments passed to evSSCM2evShape if pdim=TRUE.

Details

The spatial sign correlation is a highly robust estimator of the correlation matrix. It is consistent under elliptical distributions for the generalized correlation matrix (derived from the shape matrix instead of the correlation matrix, i.e., it is also defined when second moments are not finite).

There are two possibilities to calculate this matrix, one can either estimate all pairwise correlations by the two-dimensional spatial sign correlation or calculate the whole matrix at once by the p-dimensional spatial sign correlation. Both approaches have advantages and disadvantages. The first method should be more robust, especially if only some components of the observations are corrupted. Furthermore the consistency transformation is explicitly known only for the bivariate spatial sign correlation, whereas one has to apply an approximation procedure for the p-dimensional one. Additional argments can be passed to this algorithm using the ... argument, see the help page of SSCM2Shape for details. On the other hand, the p-dimensional spatial sign correlation is more efficient under the normal distribution and always yields a positive semidefinite estimation.

The correlation estimator is computed in three steps: the data is standardized marginally, i.e., each variable is divided by a scale estimate. (This step is optional, but recommended, and hence the default.) Then, if pdim=FALSE, for each pair of variables the 2x2 spatial sign covariance matrix (SSCM) is computed, and then from the SSCM a univariate correlation estimate given by the formulas (5) and (6) in Dürre et al. (2015). These pairwise correlation estimates are the off-diagonal elements of the returned matrix estimate. Otherwise, if pdim=TRUE, the pxp SSCM is computed, and then from the SSCM an estimator of the correlation matrix, which is done by the function SSCM2Shape, see there for details.

Scale estimation:

The scale estimates may either be computed outside the function sscor and passed on to sscor as a p-variate numeric vector, or they may be computed by sscor, using one of the following options:

"mad": applies mad from the standard package stats. This is the default.

"Qn": applies Qn from the package robustbase.

"sd": applies the standard deviation sd.

Standardizing the data is recommended (and is hence done by default), particularly so if the marginal scales largely differ. In this case, estimation without prior marginal standardization may become inefficient.

Location estimation:

The SSCM requires a multivariate location estimate. The location may be computed outside the function sscor and the result passed on to sscor as a p-variate numeric vector. Alternatively it may be computed by sscor, using one of the following options:

"2dim-median": two-dimensional spatial median, individually for every 2x2 SSCM. This is the default if pdim=FALSE.

"1dim-median": the usual, one-dimensional median applied component-wise.

"pdim-median": the p-dimensional spatial median for all variables. This is the default if pdim=TRUE.

"mean": the p-dimensional mean. In light of robustness, it is not recommended to use the mean.

There is no handling of missing values.

Value

p x p symmetric numerical matrix, the diagonal entries are 1, the off-diagonal entries are the pairwise spatial sign correlation estimates.

References

Dürre, A., Vogel, D., Fried, R. (2015): Spatial sign correlation, *Journal of Multivariate Analyis*, vol. 135, 89–105. arvix 1403.7635

Dürre, A., Vogel, D. (2016): Asymptotics of the two-stage spatial sign correlation, *Journal of Multivariate Analyis*, vol. 144, 54–67. arxiv 1506.02578

See Also

Ordinary, non-robust correlation matrix: cor.

A number of other robust correlation estimators are provided by the package rrcov.

Testing for spatial sign correlation: sscor.test.

Examples

```
set.seed(5)
X <- cbind(rnorm(25),rnorm(25))
# X is a 25x2 matrix

# cor() and sscor() behave similar under normality
sscor(X)
cor(X)

# but behave differently in the presence of outliers.
X[1,] <- c(10,10)
sscor(X)
cor(X)
```

---

sscor.test   *Correlation test based on spatial signs*

Description

Robust one-sample test and confidence interval for the correlation coefficient.

Usage

```
sscor.test(x, y, rho0=0, alternative=c("two.sided","less","greater"),
conf.level=0.95, ...)
```

Arguments

x, y	(required) numeric vectors of observations, must have the same length.
rho0	(optional) correlation coefficient under the null hypothesis. The default is 0.
alternative	(optional) character string indicating the type of alternative to be tested. Must be one of "two.sided", "less", "greater". The default is "two-sided".
conf.level	(optional) confidence level. The default is 0.95.
...	optional arguments passed to sscor (such as location and scale estimates to be used).

Details

The test is based on the spatial sign correlation (Dürre et al. 2015), which is a highly robust correlation estimator, consistent for the generalized correlation coefficient under ellipticity. The confidence interval and the p-value are based on the asymptotic distribution after a variance-stabilizing transformation similar to Fisher's z-transform. They provide accurate approximations also for very small samples (Dürre and Vogel, 2015). The test is furthermore distribution-free within the elliptical model. It has, e.g., the same power at the elliptical Cauchy distribution as at the multivariate Gaussian distribution.

Value

A list with class `"htest"` containing the following values (similar to the output of `cor.test`):

`statistic`	the value of the test statistic. Under the null, the test statistic is (asymptotically) standard normal.
`p.value`	the p-value of the test.
`estimate`	the estimated spatial sign correlation.
`null.value`	the true correlation under the null hypothesis.
`alternative`	a character string describing the alternative hypothesis.
`method`	a characters string indicating the choosen correlation estimator. Currently only the spatial sign correlation is implemented.
`data.name`	a character giving the names of the data.
`conf.int`	confidence interval for the correlation coefficient.

References

Dürre, A., Vogel, D., Fried, R. (2015): Spatial sign correlation, *Journal of Multivariate Analyis*, vol. 135, 89–105. arvix 1403.7635

Dürre, A., Vogel, D. (2015): Asymptotics of the two-stage spatial sign correlation, preprint. arxiv 1506.02578

See Also

Classical correlation testing: `cor.test`.

For more information on the spatial sign correlation: `sscor`.

Examples

```
set.seed(5)
require(mvtnorm)

# create bivariate shape matrix with correlation 0.5
sigma <- matrix(c(1,0.5,0.5,1),ncol=2)

# under normality, both tests behave similarly
data <- rmvnorm(100,c(0,0),sigma)
x <- data[,1]
y <- data[,2]

sscor.test(x,y)
cor.test(x,y)

# sscor.test also works at a Cauchy distribution
data <- rmvt(100,diag(1,2), df=1)
x <- data[,1]
y <- data[,2]

sscor.test(x,y)
cor.test(x,y)
```

Index