Package ‘mcauchyd’

July 5, 2024

Title Multivariate Cauchy Distribution; Kullback-Leibler Divergence

Version 1.3.2

Maintainer Pierre Santagostini <pierre.santagostini@agrocampus-ouest.fr>

Description Distance between multivariate Cauchy distributions, as presented by N. Bouhlel and D. Rousseau (2022) <doi:10.3390/e24060838>. Manipulation of multivariate Cauchy distributions.

License GPL (>= 3)

URL https://forgemia.inra.fr/imhorphen/mcauchyd

BugReports https://forgemia.inra.fr/imhorphen/mcauchyd/-/issues

Encoding UTF-8

RoxygenNote 7.3.2

Imports rgl, MASS, lifecycle, data.table

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Pierre Santagostini [aut, cre],
Nizar Bouhlel [aut]

Repository CRAN

Date/Publication 2024-07-05 14:30:02 UTC

Contents

mcauchyd-package *Tools for Multivariate Cauchy Distributions*

Description

This package provides tools for multivariate Cauchy distributions (MCD):

- Calculation of distances/divergences between MCD:
  - Kullback-Leibler divergence: `kldcauchy`
- Tools for MCD:
  - Probability density: `dmcd`
  - Simulation from a MCD: `rmcd`
  - Plot of the density of a MCD with 2 variables: `plotmcd`, `contourmcd`

Author(s)

Pierre Santagostini pierre.santagostini@agrocampus-ouest.fr, Nizar Bouhlel nizar.bouhlel@agrocampus-ouest.fr

References

N. Bouhlel, D. Rousseau, A Generic Formula and Some Special Cases for the Kullback–Leibler Divergence between Central Multivariate Cauchy Distributions. Entropy, 24, 838, July 2022. doi:10.3390/e24060838 #' @keywords internal

See Also

Useful links:

- `https://forgemia.inra.fr/imhorphen/mcauchyd`
- Report bugs at `https://forgemia.inra.fr/imhorphen/mcauchyd/-/issues`

contourmcd *Contour Plot of the Bivariate Cauchy Density*

Description

Draws the contour plot of the probability density of the multivariate Cauchy distribution with 2 variables with location parameter `mu` and scatter matrix `Sigma`.

Usage

```
contourmcd(mu, Sigma,
                        xlim = c(mu[1] + c(-10, 10)*Sigma[1, 1]),
                        ylim = c(mu[2] + c(-10, 10)*Sigma[2, 2]),
                        zlim = NULL, npt = 30, nx = npt, ny = npt,
                        main = "Multivariate Cauchy density",
                        sub = NULL, nlevels = 10,
                        levels = pretty(zlim, nlevels), tol = 1e-6, ...)
```

Arguments

`mu`	length 2 numeric vector.
`Sigma`	symmetric, positive-definite square matrix of order 2. The scatter matrix.
`xlim, ylim`	x-and y- limits.
`zlim`	z- limits. If NULL, it is the range of the values of the density on the x and y values within `xlim` and `ylim`.
`npt`	number of points for the discretisation.
`nx, ny`	number of points for the discretisation among the x- and y- axes.
`main, sub`	main and sub title, as for `title`.
`nlevels, levels`	arguments to be passed to the `contour` function.
`tol`	tolerance (relative to largest variance) for numerical lack of positive-definiteness in Sigma, for the estimation of the density. see `dmcd`.
`...`	additional arguments to `plot.window`, `title`, `Axis` and `box`, typically graphical parameters such as `cex.axis`.

Value

Returns invisibly the probability density function.

Author(s)

Pierre Santagostini, Nizar Bouhlel

References

N. Bouhlel, D. Rousseau, A Generic Formula and Some Special Cases for the Kullback–Leibler Divergence between Central Multivariate Cauchy Distributions. Entropy, 24, 838, July 2022. doi:10.3390/e24060838

See Also

`dmcd`: probability density of a multivariate Cauchy density

`plotmcd`: 3D plot of a bivariate Cauchy density.

Examples

```
mu <- c(1, 4)
Sigma <- matrix(c(0.8, 0.2, 0.2, 0.2), nrow = 2)
contourmcd(mu, Sigma)
```

`dmcd` *Density of a Multivariate Cauchy Distribution*

Description

Density of the multivariate (p variables) Cauchy distribution (MCD) with location parameter `mu` and scatter matrix `Sigma`.

Usage

```
dmcd(x, mu, Sigma, tol = 1e-6)
```

Arguments

`x`	length p numeric vector.
`mu`	length p numeric vector. The location parameter.
`Sigma`	symmetric, positive-definite square matrix of order p. The scatter matrix.
`tol`	tolerance (relative to largest eigenvalue) for numerical lack of positive-definiteness in Sigma.

Details

The density function of a multivariate Cauchy distribution is given by:

$$f(\mathbf{x}|\boldsymbol{\mu}, \Sigma) = \frac{\Gamma\left(\frac{1+p}{2}\right)}{\pi^{p/2}\Gamma\left(\frac{1}{2}\right)|\Sigma|^{\frac{1}{2}}\left[1 + (\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{\frac{1+p}{2}}}$$

Value

The value of the density.

Author(s)

Pierre Santagostini, Nizar Bouhlel

See Also

`rmcd`: random generation from a MCD.

`plotmcd`, `contourmcd`: plot of a bivariate Cauchy density.

Examples

```
mu <- c(0, 1, 4)
sigma <- matrix(c(1, 0.6, 0.2, 0.6, 1, 0.3, 0.2, 0.3, 1), nrow = 3)
dmcd(c(0, 1, 4), mu, sigma)
dmcd(c(1, 2, 3), mu, sigma)
```

`kldcauchy` *Kullback-Leibler Divergence between Centered Multivariate Cauchy Distributions*

Description

Computes the Kullback-Leibler divergence between two random vectors distributed according to multivariate Cauchy distributions (MCD) with zero location vector.

Usage

```
kldcauchy(Sigma1, Sigma2, eps = 1e-06)
```

Arguments

`Sigma1`	symmetric, positive-definite matrix. The scatter matrix of the first distribution.
`Sigma2`	symmetric, positive-definite matrix. The scatter matrix of the second distribution.
`eps`	numeric. Precision for the computation of the partial derivative of the Lauricella D-hypergeometric function (see Details). Default: 1e-06.

Details

Given X_1, a random vector of $\mathbb{R}^p$ distributed according to the MCD with parameters $(0, \Sigma_1)$ and X_2, a random vector of $\mathbb{R}^p$ distributed according to the MCD with parameters $(0, \Sigma_2)$.

Let $\lambda_1, \dots, \lambda_p$ the eigenvalues of the square matrix $\Sigma_1 \Sigma_2^{-1}$ sorted in increasing order:

$$\lambda_1 < \cdots < \lambda_{p-1} < \lambda_p$$

Depending on the values of these eigenvalues, the computation of the Kullback-Leibler divergence of X_1 from X_2 is given by:

- if $\lambda_1 < 1$ and $\lambda_p > 1$:
  $$KL(X_1||X_2) = -\frac{1}{2} \ln \prod_{i=1}^{p} \lambda_i + \frac{1+p}{2} \Bigg(\ln \lambda_p - \frac{\partial}{\partial a} \Bigg\{ F_D^{(p)} \Bigg(a, \underbrace{\frac{1}{2}, \dots, \frac{1}{2}, a + \frac{1}{2}}_{p}; a + \frac{1+p}{2}; 1 - \frac{\lambda_1}{\lambda_p}, \dots, 1 - \frac{\lambda_{p-1}}{\lambda_p}, 1 - \frac{1}{\lambda_p} \Bigg) \Bigg\} \Bigg|_{a=0} \Bigg)$$

- if $\lambda_p < 1$:
$$KL(X_1||X_2) = -\frac{1}{2}\ln\prod_{i=1}^{p}\lambda_i - \frac{1+p}{2}\frac{\partial}{\partial a}\left\{F_D^{(p)}\left(a, \underbrace{\frac{1}{2},\dots,\frac{1}{2}}_{p}; a+\frac{1+p}{2}; 1-\lambda_1,\dots,1-\lambda_p\right)\right\}\Big|_{a=0}$$

- if $\lambda_1 > 1$:
$$KL(X_1||X_2) = -\frac{1}{2}\ln\prod_{i=1}^{p}\lambda_i - \frac{1+p}{2}\prod_{i=1}^{p}\frac{1}{\sqrt{\lambda_i}}$$
$$\times\frac{\partial}{\partial a}\left\{F_D^{(p)}\left(\frac{1+p}{2}, \underbrace{\frac{1}{2},\dots,\frac{1}{2}}_{p}; a+\frac{1+p}{2}; 1-\frac{1}{\lambda_1},\dots,1-\frac{1}{\lambda_p}\right)\right\}\Big|_{a=0}$$

where $F_D^{(p)}$ is the Lauricella D-hypergeometric function defined for p variables:

$$F_D^{(p)}\left(a; b_1, ..., b_p; g; x_1, ..., x_p\right) = \sum_{m_1 \geq 0} ... \sum_{m_p \geq 0} \frac{(a)_{m_1+...+m_p}(b_1)_{m_1} ... (b_p)_{m_p}}{(g)_{m_1+...+m_p}} \frac{x_1^{m_1}}{m_1!} ... \frac{x_p^{m_p}}{m_p!}$$

Value

A numeric value: the Kullback-Leibler divergence between the two distributions, with two attributes `attr(, "epsilon")` (precision of the partial derivative of the Lauricella D-hypergeometric function,see Details) and `attr(, "k")` (number of iterations).

Author(s)

Pierre Santagostini, Nizar Bouhlel

References

N. Bouhlel, D. Rousseau, A Generic Formula and Some Special Cases for the Kullback–Leibler Divergence between Central Multivariate Cauchy Distributions. Entropy, 24, 838, July 2022. doi:10.3390/e24060838

Examples

```
Sigma1 <- matrix(c(1, 0.6, 0.2, 0.6, 1, 0.3, 0.2, 0.3, 1), nrow = 3)
Sigma2 <- matrix(c(1, 0.3, 0.1, 0.3, 1, 0.4, 0.1, 0.4, 1), nrow = 3)
kldcauchy(Sigma1, Sigma2)
kldcauchy(Sigma2, Sigma1)

Sigma1 <- matrix(c(0.5, 0, 0, 0, 0.4, 0, 0, 0, 0.3), nrow = 3)
Sigma2 <- diag(1, 3)
# Case when all eigenvalues of Sigma1 %*% solve(Sigma2) are < 1
kldcauchy(Sigma1, Sigma2)
# Case when all eigenvalues of Sigma1 %*% solve(Sigma2) are > 1
kldcauchy(Sigma2, Sigma1)
```

`lnpochhammer` *Logarithm of the Pochhammer Symbol*

Description

Computes the logarithm of the Pochhammer symbol.

Usage

```
lnpochhammer(x, n)
```

Arguments

`x`	numeric.
`n`	positive integer.

Details

The Pochhammer symbol is given by:

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)} = x(x+1)...(x+n-1)$$

So, if $n > 0$:

$$log\left((x)_n\right) = log(x) + log(x+1) + ... + log(x+n-1)$$

If $n = 0$, $log\left((x)_n\right) = log(1) = 0$

Value

Numeric value. The logarithm of the Pochhammer symbol.

Author(s)

Pierre Santagostini, Nizar Bouhlel

See Also

`pochhammer()`

Examples

```
lnpochhammer(2, 0)
lnpochhammer(2, 1)
lnpochhammer(2, 3)
```

plotmcd *Plot of the Bivariate Cauchy Density*

Description

Plots the probability density of the multivariate Cauchy distribution with 2 variables with location parameter `mu` and scatter matrix `Sigma`.

Usage

```
plotmcd(mu, Sigma, xlim = c(mu[1] + c(-10, 10)*Sigma[1, 1]),
                ylim = c(mu[2] + c(-10, 10)*Sigma[2, 2]), n = 101,
                xvals = NULL, yvals = NULL, xlab = "x", ylab = "y",
                zlab = "f(x,y)", col = "gray", tol = 1e-6, ...)
```

Arguments

`mu`	length 2 numeric vector.
`Sigma`	symmetric, positive-definite square matrix of order 2. The scatter matrix.
`xlim`, `ylim`	x-and y- limits.
`n`	A one or two element vector giving the number of steps in the x and y grid, passed to `plot3d.function`.
`xvals`, `yvals`	The values at which to evaluate `x` and `y`. If used, `xlim` and/or `ylim` are ignored.
`xlab`, `ylab`, `zlab`	The axis labels.
`col`	The color to use for the plot. See `plot3d.function`.
`tol`	tolerance (relative to largest variance) for numerical lack of positive-definiteness in Sigma, for the estimation of the density. see `dmcd`.
`...`	Additional arguments to pass to `plot3d.function`.

Value

Returns invisibly the probability density function.

Author(s)

Pierre Santagostini, Nizar Bouhlel

References

N. Bouhlel, D. Rousseau, A Generic Formula and Some Special Cases for the Kullback–Leibler Divergence between Central Multivariate Cauchy Distributions. Entropy, 24, 838, July 2022. doi:10.3390/e24060838

See Also

`dmcd`: probability density of a multivariate Cauchy density

`contourmcd`: contour plot of a bivariate Cauchy density.

`plot3d.function`: plot a function of two variables.

Examples

```
mu <- c(1, 4)
Sigma <- matrix(c(0.8, 0.2, 0.2, 0.2), nrow = 2)
plotmcd(mu, Sigma)
```

---

pochhammer *Pochhammer Symbol*

Description

Computes the Pochhammer symbol.

Usage

```
pochhammer(x, n)
```

Arguments

x	numeric.
n	positive integer.

Details

The Pochhammer symbol is given by:

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)} = x(x+1)...(x+n-1)$$

Value

Numeric value. The value of the Pochhammer symbol.

Author(s)

Pierre Santagostini, Nizar Bouhlel

Examples

```
pochhammer(2, 0)
pochhammer(2, 1)
pochhammer(2, 3)
```

rmcd *Simulate from a Multivariate Cauchy Distribution*

Description

Produces one or more samples from the multivariate (p variables) Cauchy distribution (MCD) with location parameter `mu` and scatter matrix `Sigma`.

Usage

```
rmcd(n, mu, Sigma, tol = 1e-6)
```

Arguments

`n`	integer. Number of observations.
`mu`	length p numeric vector. The location parameter.
`Sigma`	symmetric, positive-definite square matrix of order p. The scatter matrix.
`tol`	tolerance for numerical lack of positive-definiteness in Sigma (for `mvrnorm`, see Details).

Details

A sample from a MCD with parameters $\boldsymbol{\mu}$ and Σ can be generated using:

$$\mathbf{X} = \boldsymbol{\mu} + \frac{\mathbf{Y}}{\sqrt{u}}$$

where $\mathbf{Y}$ is a random vector distributed among a centered Gaussian density with covariance matrix Σ (generated using `mvrnorm`) and u is distributed among a Chi-squared distribution with 1 degree of freedom.

Value

A matrix with p columns and n rows.

Author(s)

Pierre Santagostini, Nizar Bouhlel

See Also

`dmcd`: probability density of a MCD.

Examples

```
mu <- c(0, 1, 4)
sigma <- matrix(c(1, 0.6, 0.2, 0.6, 1, 0.3, 0.2, 0.3, 1), nrow = 3)
x <- rmcd(100, mu, sigma)
x
apply(x, 2, median)
```

Index