Package 'MKdepth'

May 26, 2020

Title Compute Monge-Kantorovich depth

Version 0.1.0

Author Meurice Leo <leomeurice@hotmail.com>

Maintainer Meurice Leo <leomeurice@hotmail.com>

Description This package allows to calculate the Monge-Kantorovich depth of a dataset and also allows to perform other tasks.

License GPL-3

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.0.9000

Depends R (>= 3.6),
alphahull,
depth,
adagio,
transport,
akima,
sp,
grDevices,
graphics,
Rdpack

RdMacros Rdpack

R topics documented:

`ball_distribution` *Simulate from a Ball Distribution*

Description

Produces samples from the uniform distribution on the p-ball.

Usage

```
ball_distribution(n, p)
```

Arguments

`n`	number of samples
`p`	ball dimension

Details

The ball distribution is the distribution of Z where

$$Z = UX,$$

where U is the inverse p-th root of a uniform distribution on $[0, 1]$ and X is uniform on the (p-1)-sphere which is generated from `sphere_distribution`.

Value

`n` by `p` matrix with one sample in each row.

Author(s)

Leo Meurice

See Also

`sphere_distribution`

Examples

```
set.seed(1)
u = ball_distribution(1000,2)
plot(u)
```

banana_distribution *Simulate from a Banana Distribution*

Description

Produces samples from a banana distribution in two dimensions.

Usage

```
banana_distribution(n, height = 1, thick = 1)
```

Arguments

n	number of samples.
height	height of the banana. By default height=1
thick	thickness of the banana. By default thick=1

Details

The banana distribution is the distribution of the vector

$$(X + Rcos\phi, height * X^2 + Rsin\phi),$$

where X is uniform on $[-1, 1]$, ϕ is uniform on $[0, 2\pi]$, Z is uniform on $[0, thick]$, X, Z and ϕ are independent, and R = 0.2Z(1 + (1 - |X|)/2).

Value

n by 2 matrix with one sample in each row.

Author(s)

Leo Meurice

References

Chernozhukov V, Galichon A, Hallin M, Henry M (2017). "Monge–Kantorovich depth, quantiles, ranks and signs." *The Annals of Statistics*, **45**(1), 223–256.

Examples

```
set.seed(1)
u = banana_distribution(1000)
plot(u)

set.seed(1)
u = banana_distribution(1000,4,.5)
plot(u)
```

MKdepth *The Monge-Kantorovich Depth*

Description

Computes the Monge-Kantorovich depth of each observation from a data set.

Usage

```
MKdepth(
  data,
  trs_method = "assignment",
  depth_method = "Tukey",
  approx = FALSE,
  prec = 3
)
```

Arguments

data	n by p matrix of data where each row contains a p-variate point, w.r.t which the Monge-Kantorovich depth is to be calculated.
trs_method	Character sting which determines the transport algorithm used. By default, trs_method="assignment". See details below.
depth_method	Character string which determines the depth function used for the reference distribution. Method can be "Tukey" (the default), "Liu" or "Oja".
approx	Logical. If dimension is greater than or equal to 3, the Tukey depth can be approximate . Useful when sample size is large.
prec	Integer indicating the number of decimal places to be used for depth values. By default, prec=3.

Details

The Monge-Kantorovich depth is constructed as explained in (Chernozhukov et al. 2017).

In order to calculate the Monge-Kantorovich depth, we need to calculate the Tukey depth of a specific distribution. For this, we use the function depth of the package depth. For more details on the depth method, please see depth details.

The parameter trs_method is the transport method which is used for the Monge-Kantoroch depth. The default method are: "assignement", comes from the adagio package (See References). The other methods which can be used, come from the transport package and are the same as for the transport function (See References).

Value

Numerical vector of depths, one for each row in data.

Author(s)

Leo Meurice

References

Genest M, Masse J, contains eigen JPs, tql2, written by the EISPLACK authors t, dgedi, dgefa from LINPACK, daxpy, dscal, dswap, idamax from LINPACK, from NAPACK V, written by J. C. Gower A7, written by F. K. Bedall A1, Zimmermann H, written by P.J. Rousseeuw A3, Ruts. I (2019). *depth: Nonparametric Depth Functions for Multivariate Analysis*. R package version 2.1-1.1, https://CRAN.R-project.org/package=depth.

Borchers HW (2018). *adagio: Discrete and Global Optimization Routines*. R package version 0.7.1, https://CRAN.R-project.org/package=adagio.

Schuhmacher D, Bahre B, Gottschlich C, Hartmann V, Heinemann F, Schmitzer B (2019). *transport: Computation of Optimal Transport Plans and Wasserstein Distances*. R package version 0.12-1, https://cran.r-project.org/package=transport.

Chernozhukov V, Galichon A, Hallin M, Henry M (2017). "Monge–Kantorovich depth, quantiles, ranks and signs." *The Annals of Statistics*, **45**(1), 223–256.

Examples

```
set.seed(1)
data=ball_distribution(1000,2)
depth=MKdepth(data)

set.seed(1)
data=matrix(rnorm(2000),1000)
depth=MKdepth(data,trs_method="auction")
```

`MKdepth.package` *MKdepth Package*

Description

This package contains functions allowing to calculate the depth of Monge-Kantorovich and allowing to perform other similar tasks like drawing depth contours.

Author(s)

Leo Meurice <leomeurice@hotmail.com>

See Also

`ball_distribution` `banana_distribution` `MKdepth` `MKdepth_contours` `MKdepth_iterative` `MKdepth_map` `MKdepth_median` `MKdepth_specific` `sphere_distribution`

`MKdepth_contours` *Monge-Kantorovich Depth Contours*

Description

Builds the Monge-Kantorovich depth contours for 2-dimensional data.

Usage

```
MKdepth_contours(
  data,
  depth,
  alpha,
  contours = 3,
  col = topo.colors(contours),
  legend = F,
  plot.new = T
)
```

Arguments

`data`	`n` by `2` matrix of data where each row contains a two dimensional point, w.r.t which the Monge-Kantorovich depth is calculated. `n` must be larger than 20.
`depth`	Numerical vector of depths, one for each row in `data`.
`alpha`	Value of α. This parameter is defined in the function `ahull` from the `alphahull-package`. (See Details below.)
`contours`	Integer. Contours number to be drawn. By default, `contours=3`.
`col`	Color, used to draw contours. The length must be equal to `contours`.
`legend`	Logical; if `TRUE`, the legend is displayed. By default, `legend=F`.
`plot.new`	Logical; if `TRUE`, a new plot is displayed and the data points and contours are drawn.

Details

The depth contours are drawn using the `ahull` function from the `alphahull-package`. The `alpha` parameter must be adjusted to conform to the shape of the non-convex depth contours.

Author(s)

Leo Meurice

References

Pateiro-Lopez B, Rodriguez-Casal A (2019). *alphahull: Generalization of the Convex Hull of a Sample of Points in the Plane*. R package version 2.2, `https://CRAN.R-project.org/package=alphahull`.

Examples

```
set.seed(1)
data=ball_distribution(100,2)
depth=MKdepth_iterative(data,iter=100)
MKdepth_contours(data,depth$MKdepth,alpha=1)

set.seed(1)
data=matrix(rnorm(2000),1000)
depth=MKdepth(data)
plot(data,pch='.')
MKdepth_contours(data,depth,contours=5,alpha=2,plot.new=FALSE)
```

`MKdepth_iterative` *The Monge-Kantorovich Depth Iterative Function*

Description

Calculates the Monge-Kantorovich depth using several iterations in order to obtain more reliable values. The function `MKdepth` is computed several times to return the median of the different results.

Usage

```
MKdepth_iterative(
  data,
  trs_method = "assignment",
  depth_method = "Tukey",
  approx = FALSE,
  iter = 1,
  eps = NULL,
  prec = 3
)
```

Arguments

`data`	n by p matrix of data where each row contains a p-variate point, w.r.t which the Monge-Kantorovich depth is to be calculated.
`trs_method`	Character sting which determines the transport algorithm used. By default, `trs_method="assignment"`. See details below.
`depth_method`	Character string which determines the depth function used for the reference distribution. Method can be "Tukey" (the default), "Liu" or "Oja".
`approx`	Logical. If dimension is greater than or equal to 3, the Tukey depth can be approximate . Useful when sample size is large.
`iter`	Positive integer. Iterations number, i.e , the calls number of the function `MKdepth`. If `eps` is specified, `iter` is not considered. By default `iter=1`.
`eps`	Positive double. Maximum expected difference between the depth values of the penultimate iteration and the depth values of the last iteration. If `eps=NULL`, this is `iter`, the iterations number, which is considered. By default `eps=NULL`.
`prec`	Integer indicating the number of decimal places to be used for depth values. By default, `prec=3`.

Details

Due to the random generate of a ball-distribution, the computed depth values by the MKdepth function may not be precisely exact. In order to counter this, the function MKdepth is computed several times and the median of all results is selected. To limit the number of iterations, the parameters iter and eps are used.

Value

List containing a numerical vector of depths ($MKdepth), one for each row in data, the iterations number ($iteration) and the execution time ($time).

Author(s)

Leo Meurice

Examples

```
data=ball_distribution(100,2)
depth=MKdepth_iterative(data,eps=0.01)

data=matrix(rnorm(2000),1000)
depth=MKdepth_iterative(data,iter=10)
```

MKdepth_map *Level plot for Monge-Kantorovich Depth*

Description

Produces a contour plot with the areas between the contours filled in solid color. A key showing how the colors map to depth values is shown to the right of the plot. Construct from filled.contour fonction.

Usage

```
MKdepth_map(
  data,
  depth,
  alpha,
  by,
  col = heat.colors,
  main = "",
  xlab = "",
  ylab = ""
)
```

Arguments

data	n by 2 matrix of data where each row contains a two dimensional point, w.r.t which the Monge-Kantorovich depth is calculated.
depth	Numerical vector of depths, one for each row in data.

alpha	Value of α. This parameter is defined in the function ahull from the alphahull-package. (See Details below.)
by	Interpolation step.
col	color, used to draw level plot.
main	an overall title for the plot (optional).
xlab	a title for the x axis (optional).
ylab	a title for the y axis (optional).

Author(s)

Leo Meurice

References

Pateiro-Lopez B, Rodriguez-Casal A (2019). *alphahull: Generalization of the Convex Hull of a Sample of Points in the Plane*. R package version 2.2, https://CRAN.R-project.org/package=alphahull.

Examples

```
set.seed(1)
data=ball_distribution(1000,2)
depth=MKdepth(data)
MKdepth_map(data,depth,alpha=.5,by=.1)

set.seed(1)
data=matrix(rnorm(2000),1000)
depth=MKdepth(data)
MKdepth_map(data,depth,alpha=2,by=.1)
```

MKdepth_median *Monge-Kantorovich Median*

Description

Computes the Monge-Kantorovich Median of a multivariate data set.

Usage

```
MKdepth_median(data, depth)
```

Arguments

data	n by p matrix of data where each row contains a p-variate point, w.r.t which the Monge-Kantorovich depth is calculated.
depth	Numerical vector of depths, one for each row in data.

Value

List containing the Monge-Kantorovich median coordinates($coordinates) and its depth value($depth).

Author(s)

Leo Meurice

Examples

```
set.seed(1)
data=ball_distribution(1000,2)
depth=MKdepth(data)
median=MKdepth_median(data,depth)
plot(data,pch='.')
points(median$coordinates[1],median$coordinates[2],pch=19,col='red')

set.seed(1)
data=matrix(rnorm(2000),1000)
depth=MKdepth(data)
median=MKdepth_median(data,depth)
plot(data,pch='.')
MKdepth_contours(data,depth,contours=5,alpha=2,plot.new=FALSE)
points(median$coordinates[1],median$coordinates[2],pch=19,col='red')
```

`MKdepth_specific` *The Monge-Kantorovich Depth of specific observations compared a data set*

Description

Calculates the Monge-Kantorovich depth of each specific observation of a data set.

Usage

```
MKdepth_specific(
  x,
  data,
  trs_method = "assignment",
  depth_method = "Tukey",
  approx = FALSE,
  iter = 1,
  eps = NULL,
  prec = 3
)
```

Arguments

`x`	n by p matrix of data whose Monge-Kantorovich depth is to be calculated. Each row contains a p-variate point. Should have the same dimension as `data`.
`data`	n by p matrix of data where each row contains a p-variate point, w.r.t which the Monge-Kantorovich depth is to be calculated.
`trs_method`	Character sting which determines the transport algorithm used. By default, `trs_method="assignment"`. See details below.
`depth_method`	Character string which determines the depth function used for the reference distribution. Method can be "Tukey" (the default), "Liu" or "Oja".

approx	Logical. If dimension is greater than or equal to 3, the Tukey depth can be approximate . Useful when sample size is large.
iter	Positive integer. Iterations number, i.e , the calls number of the function MKdepth. If eps is specified, iter is not considered. By default iter=1.
eps	Positive double. Maximum expected difference between the depth values of the penultimate iteration and the depth values of the last iteration. If eps=NULL, this is iter, the iterations number, which is considered. By default eps=NULL.
prec	Integer indicating the number of decimal places to be used for depth values. By default, prec=3.

Details

Call the function MKdepth_iterative.

Value

Numerical vector of depths, one for each row in x.

Author(s)

Leo Meurice

See Also

MKdepth_iterative

Examples

```
set.seed(1)
data=ball_distribution(100,2)
x=rbind(c(-0.2,0.1),c(-.25,-.5),c(.5,0))
depth=MKdepth_specific(x,data,iter=20)
plot(data,pch=20)
points(x,pch=19,col='red')
text(x[,1], x[,2], depth,cex=1, pos=3,col="red")
```

---

sphere_distribution *Simulate from a Sphere Distribution*

Description

Produces samples from the uniform distribution on the p-sphere.

Usage

```
sphere_distribution(n, p)
```

Arguments

n	number of samples
p	sphere dimension

Details

The sphere distribution is the distribution of Y where

$$Y = \frac{X}{\|X\|},$$

where X follows a multinomial distribution.

Value

n by p matrix with one sample in each row.

Author(s)

Leo Meurice

Examples

```
u = sphere_distribution(100,2)
plot(u)
```

Index