Package 'iccTraj'

November 2, 2023

Type Package

Title Estimates the Intraclass Correlation Coefficient for Trajectory Data

Version 1.0.4

Depends R (>= 4.0)

Imports doParallel, dplyr, magic, trajectories, sp, spacetime, purrr, utils, foreach

Description Estimates the intraclass correlation coefficient for trajectory data using a matrix of distances between trajectories. The distances implemented are the extended Hausdorff distances (Min et al. 2007) <doi:10.1080/13658810601073315> and the discrete Fréchet distance (Magdy et al. 2015) <doi:10.1109/IntelCIS.2015.7397286>.

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

NeedsCompilation no

Author Josep L. Carrasco [aut, cre]

Maintainer Josep L. Carrasco <jlcarrasco@ub.edu>

Repository CRAN

Date/Publication 2023-11-02 08:20:05 UTC

R topics documented:

gull_data 2
HD 2
ICC 3
iccTraj 4
interval 6

Index **7**

gull_data *Gull data*

Description

A data frame with sample of 90 gull trajectories.

Usage

```
gull_data
```

Format

A data frame containing 90 trajectories

ID Subject identifier

trip Trip identifier

LONG Longitude

LAT Latitude

triptime Time in seconds when the locations were obtained

HD *Computes extended Hausdorff distance between two trajectories.*

Description

Computes extended Hausdorff distance between two trajectories.

Usage

```
HD(pp1, pp2, q = 1)
```

Arguments

pp1	Set of spatial points for the first trajectory. It can be a matrix of 2D points, first column x/longitude, second column y/latitude, or a SpatialPoints or SpatialPointsDataFrame object.
pp2	Set of spatial points for the second trajectory. It can be a matrix of 2D points, first column x/longitude, second column y/latitude, or a SpatialPoints or SpatialPointsDataFrame object.
q	Quantile for the extended Hausdorff distance. Default value q=1 uses the maximum that leads to classical Hausdorff distance.

Value

A numerical value with the distance.

References

Magdy, N., Sakr, M., Abdelkader, T., Elbahnasy, K. (2015). Review on trajectory similarity measures. 10.1109/IntelCIS.2015.7397286.

Min, D., Zhilin, L., Xiaoyong, C. (2007) Extended Hausdorff distance for spatial objects in GIS. International Journal of Geographical Information Science, 21:4, 459–475

Examples

```
# Take two trajectories
library(dplyr)
library(sp)
sample_data<-gull_data %>% filter(ID %in% c(5107912,5107913), trip %in% c("V02","V01"))
tr1<-gull_data %>% filter((ID == 5107912) & (trip=="V02"))
tr2<-gull_data %>% filter((ID == 5107913) & (trip=="V01"))
pts1 = SpatialPoints(tr1[c("LONG","LAT")], proj4string=CRS("+proj=longlat"))
pts2 = SpatialPoints(tr2[c("LONG","LAT")], proj4string=CRS("+proj=longlat"))
# Hausdorff distance
HD(pts1,pts2,q=1)
# Median Hausdorff distance
HD(pts1,pts2,q=0.5)
```

ICC *Computes the intraclass correlation coefficient (ICC) using a matrix of distances.*

Description

Computes the intraclass correlation coefficient (ICC) using a matrix of distances.

Usage

```
ICC(X, nt)
```

Arguments

X	Matrix with the pairwise distances.
nt	Data frame with the number of trips by subject

Details

The intraclass correlation coeffcient is estimated using the distance matrix among trajectories.

Value

Data frame with the estimates of the ICC (r), the subjects' mean sum-of-squares (MSA), the between-subjects variance (sb), the total variance (st), and the within-subjects variance (se).

`iccTraj` *Estimates the intraclass correlation coefficient (ICC) for trajectory data*

Description

Estimates the intraclass correlation coefficient (ICC) for trajectory data

Usage

```
iccTraj(
  data,
  ID,
  trip,
  LON,
  LAT,
  time,
  projection = CRS("+proj=longlat"),
  origin = "1970-01-01 UTC",
  parallel = TRUE,
  individual = TRUE,
  distance = c("H", "F"),
  bootCI = TRUE,
  nBoot = 100,
  q = 0.5
)
```

Arguments

`data`	A data frame with the locations and times of trajectories. It is assumed the time between locations is uniform. It must contain at least five columns: subject identifier, trip identifier, latitude, longitude, and time of the reading.
`ID`	Character string indicating the name of the subjects column in the dataset.
`trip`	Character string indicating the trip column in the dataset.
`LON`	Numeric. Longitude readings.
`LAT`	Numeric. Latitude readings.
`time`	Numeric. Time of the readings.
`projection`	Projection string of class CRS-class.
`origin`	Optional. Origin of the date-time. Only needed in the internal process to create an object of type POSIXct.

parallel	TRUE/FALSE value. Use parallel computation? Default value is TRUE.
individual	TRUE/FALSE value. Compute individual within-subjects variances? Default value is TRUE.
distance	Metric used to compute the distances between trajectories. Options are **H** for median Hausforff distance, and **F** for discrete Fréchet distance.
bootCI	TRUE/FALSE value. If TRUE it will generate boostrap resamples. Default value is TRUE.
nBoot	Numeric. Number of bootstrap resamples. Ignored if "bootCI" is FALSE. Default value is 100.
q	Quantile for the extended Hausdorff distance. Default value q=0.5 leads to median Hausdorff distance.

Details

The intraclass correlation coefficient is estimated using the distance matrix among trajectories.

Bootstrap resamples are obtained using balanced randomized cluster bootstrap approach (Davison and Hinkley, 1997; Field and Welsh, 2007)

Value

An object of class *iccTraj*.The output is a list with the following components:

- *est*. Data frame with the following estimates: the ICC (r), the subjects' mean sum-of-squares (MSA), the between-subjects variance (sb), the total variance (st), and the within-subjects variance (se).
- *boot*. If bootCI argument is set to TRUE, data frame with the bootstrap estimates.
- *D*. Data frame with the pairwise distances among trajectories.
- *indW* Data frame with the following columns: the subject's identifier (ID), the individual within-subjects variances (w), the individual ICC (r), and the number of trips (n).

References

Davison A.C., Hinkley D.V. (1997). Bootstrap Methods and Their Application. Cambridge: Cambridge University Press.

Field, C.A., Welsh, A.H. (2007). Bootstrapping Clustered Data. Journal of the Royal Statistical Society. Series B (Statistical Methodology). 69(3), 369-390.

Examples

```
# Using median Hausdorff distance.
 Hd<-iccTraj(gull_data,"ID","trip","LONG","LAT","triptime")
 Hd$est
# Using discrete Fréchet distance.
Fd<-iccTraj(gull_data,"ID","trip","LONG","LAT","triptime", distance="F")
Fd$est
```

`interval` *Computes the confidence interval for the ICC*

Description

Computes the confidence interval for the ICC

Usage

```
interval(x, conf = 0.95, method = c("EB", "AN", "ZT"))
```

Arguments

`x`	An object of class `"iccTraj"`
`conf`	Numeric. Level of confidence. Default is set to 0.95.
`method`	String. Method used to estimate the confidence interval. Accepted values are **EB** for Empirical Bootstrap, **AN** for asymptotic Normal, and **ZT** for asymptotic Normal using the Z-transformation.

Details

Let $\hat{\theta}$ denote the ICC sample estimate and θ_i^B denote the ICC bootstrap estimates with $i = 1, \ldots, B$. Let $\delta^B_{\alpha/2}$ and $\delta^B_{1-\alpha/2}$ be the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ percentiles of $\delta_i^B = \theta_i^B - \hat{\theta}$. The empirical bootstrap confidence interval is then estimated as $\hat{\theta} + \delta^B_{\alpha/2}, \hat{\theta} + \delta^B_{1-\alpha/2}$.

Asymptotic Normal (AN) interval is obtained as $\hat{\theta} \pm Z_{1-\alpha/2} * SE_B$ where SE_B denotes the standard deviation of θ_i^B, and $Z_{1-\alpha/2}$ stands for the $1 - \alpha/2$ quantile of the standard Normal distribution.

In the ZT approach, the ICC is transformed using Fisher's Z-transformation. Then, the AN approach is applied to the transformed ICC.

Value

A vector with the two boundaries of the confidence interval.

Examples

```
# Using median Hausdorff distance
Hd<-iccTraj(gull_data,"ID","trip","LONG","LAT","triptime", parallel=FALSE, distance="H")
Hd$est
interval(Hd)
```

Index

∗ **datasets**
gull_data, 2

gull_data, 2

HD, 2

ICC, 3
iccTraj, 4
interval, 6