Package 'AST'

August 16, 2018

Type Package

Title Age-Spatial-Temporal Model

Version 0.1.0

Author Parinaz Mehdipour <mehdipoor.p@gmail.com> [aut], Ali Ghanbari <a.ghanbari541@gmail.com> [cre,aut], Farshad Farzadfar <farzadfar3@yahoo.com> [cph]

Maintainer Ali Ghanbari <a.ghanbari541@gmail.com>

Description Fits a model to adjust and consider additional variations in three dimensions of age groups, time, and space on residuals excluded from a prediction model that have residual such as: linear regression, mixed model and so on. Details are given in Foreman et al. (2015) <doi:10.1186/1478-7954-10-1>.

License GPL-2 | GPL-3

Encoding UTF-8

LazyData true

Depends R (>= 3.0.0), stats

NeedsCompilation no

Repository CRAN

Date/Publication 2018-08-16 10:30:17 UTC

R topics documented:

1

AST	*Age-Spatio-Temporal Model*

Description

The age-spatial-temporal models utilizes additional prediction analysis to take into account how the dependent variable further varies across time, space, and age. This type of age-spatial-temporal model has been used in many applications. We do this by calculating the residual (predicted - observed dependent variable) for each data point and then run local regression in three dimensions on the residual. The process assumes that residuals contain valuable information that cannot be directly observed but nonetheless vary systematically across geographic region, time, and age group. This allows us to predict how much the observed dependent variable differs from the regression model's prediction and to account for these differences. This extra variation in prediction model defines by three matrices of age, time and spaces. Once these weights have been calculated, weighting every observation in the data set relative to the one being predicted, it a simple matter of calculating a weighted average of the residuals from regression. This "predicted residual" is then added back onto the prediction, creating an estimate that more closely takes into account aspects of the data that cannot be captured by a simple covariate model. This package also considers the data sources' weight in calculating. But this is optional to include.

Usage

```
AST(data.residual, spaceMatrix, par.time=0.5, par.age=1,
    weight.coverage=0.9, agecat, minyear, maxyear)
```

Arguments

data.residual	Data frame of residuals from any regression model. It should be included age, year,location and residual variables. The corresponding numeric variable names in data set are residuals of any prediction model, year, location, and age. coverage variable is also optional for data sources and data points which are more reliable and representative.
spaceMatrix	A square matrix of location which the number of rows and columns are equal to the number of spaces. Colnames and rownames must be the ordered number of locations and match with the name of location variable in data.residual.
par.time	A positive parameter which can be tuned to increase or decrease how much smoothing occurs across time. Usual range is 0.5-2.0, higher value results in less strength drawn across time.
par.age	A positive parameter which can be tuned to increase or decrease how much smoothing occurs across age groups. Usual range is 0.5-2.0, higher value results in less strength drawn across age.
weight.coverage	
	A parameter which can be specified to give greater weights to data sources that are more reliable. If data frame does not include coverage variable, this parameter will not be considered in AST function. This parameter assigns to data points that their corresponding coverage variable is 1. [1-weight.coverage] computes for 0 ones.

agecat	The number of age groups which were predicted. It sets to 1 when we do not have age variable in data frame.
minyear	The first year of data frame which is predicted.
maxyear	The last year of data frame which is predicted.

Details

This package adjusts residuals of any prediction model like regression by time, age, and spaces. The weighted residuals then added back to predictions.

Space matrix specifies the amount of weights in location. It should be a square matrix or data frame with equal number of rows and columns with location variable in data.residual. This data set is define by users with desirable amount of weights with adjacent locations or even how smooth they can be in far from areas. If users do not have this information, calcSpaceMat is a function in this package which builds a matrix of adjacent locations from a shapefile and give them equal weights.

Space matrix can be handled by downloading a shapefile. The shapefile format is a popular geospatial vector data format for geographic information system (GIS) software.The term "shapefile" is quite common, but is misleading since the format consists of a collection of files with a common filename prefix, stored in the same directory. The three mandatory files have filename extensions .shp, .shx, and .dbf. In example part, we used a package to read this file format and determine a matrix of adjacent locations.

Value

a list including three parts as given by:

1. adj.rate a data frame that includes 5 variables: ID, time, year, locations, and adjusted residuals.

2. Age_weight matrix

3. time_weight matrix

Note

The age, time, and space variables should be ordered.

Author(s)

Parinaz Mehdipour, Ali Ghanbari

References

Foreman et al.: Modeling causes of death: an integrated approach using CODEm. Population Health Metrics 2012 10:1.

Examples

```
library("AST")

data(data.residual.AST,package="AST")
data.residual <- data.residual.AST[data.residual.AST$sex == 0 ,]
```

```
data(data.adjacent.mat,package="AST")
spaceMat <- calcSpaceMat(data.adjacent.mat,par.space=0.9)

AST(data.residual=data.residual,spaceMatrix=spaceMat,par.time = 0.5,
     par.age=1, weight.coverage=0.9, agecat=3, minyear=2000, maxyear=2005)
```

---

calcSpaceMat   *Space Matrix Calculation for AST*

Description

a function that calculates the space matrix for AST model by determining adjacent locations and give them equal weights 1 and 0 for non-neighbors.

Usage

```
calcSpaceMat(adjacent.mat, par.space = 0.9)
```

Arguments

adjacent.mat   a matrix for adjacent locations with weight 1 and 0 for non-neighbors. Row-
                names of this matrix must be match with location variable.

par.space   A parameter which specify the amount of correlation between adjacent loca-
                tions.

Value

a square matrix

Note

rowname of space matrix should be numeric and the same as location variable in data.residual.

If a shapefile is used for making adjacent matrix, package "spdep" can be used to identify neighbors.

Author(s)

Parinaz Mehdipour, Ali Ghanbari

Examples

```
library("AST")
data(data.adjacent.mat,package="AST")
calcSpaceMat(data.adjacent.mat,par.space=0.9)
```

---

data.adjacent.mat *data frame of adjacent provinces in Iran*

Description

This a matrix belonging to 31 provinces of Iran. The rownames of this matrix are matched with the codes of administrative names in Iran. In this matrix adjacent provinces specify with 1 and others with 0.

Usage

```
data("data.adjacent.mat")
```

Format

A matrix of adjacent provinces with equal rows and columns

Examples

```
data(data.adjacent.mat)
```

---

data.residual.AST *residual data set*

Description

a data frame of residuals excluded from a prediction model

Usage

```
data("data.residual.AST")
```

Format

A data frame with 9512 observations on the following 6 variables.

year a numeric vector that specify the time dimension in our predictions

sex a numeric vector

age a numeric vector that specify the age dimension in our predictions. It is categorical and sorted.

residual a numeric vector of residuals from Generalized Linear Mixed Model (GLMM) that is the selected prediction model for this data set.

coverage a numeric vector that specify the amount of reliability for some data points, National data sources get coverage equal to 1 and other sub-national data sources and systematic reviews get 0. The coverage weight is applied to data sources with the amount of 1 in coverage variable.

location a numeric vector that specify the space dimension in our predictions and in this data set is Iran's provinces.

Details

This is a sub-sample of Body Mass Index (BMI) real data from National and Sub-national Burden of Diseases, Injuries, and Metabolic Risk Factors in Iran (NASBOD). There are nine different data sources for BMI that some of them are more reliable and representative and we give higher weights to them.

Source

https://www.ncdrc.info/

Examples

```
data(data.residual.AST)
```

Index