Package ‘sphet’

April 13, 2012

Version 1.1-12

Date 2012-04-04

Title Spatial Models with heteroskedastic innovations

Author Gianfranco Piras <gpiras@mac.com>

Maintainer Gianfranco Piras <gpiras@mac.com>

Depends R (>= 2.15), nlme, spdep, Matrix, fields

Description GM estimation and Spatial HAC

License GPL-2

LazyLoad yes

LazyData no

Repository CRAN

Repository/R-Forge/Project sphet

Repository/R-Forge/Revision 43

Date/Publication 2012-04-13 20:28:29

R topics documented:

`sphet-package` *Estimation of spatial models with heteroskedastic innovations*

Description

A set of functions to estimate spatial models with heteroskedastic innovations

Details

Package:	sphet
Type:	Package
Version:	1.0-0
Date:	2010-06-03
License:	GPL
LazyLoad:	yes

Author(s)

Gianfranco Piras <gpiras@mac.com>

References

Piras, Gianfranco (2010) sphet: Spatial Models with Heteroskedastic Innovations in R, *Journal of Statistical Software* June 2010, Volume 35, Issue 1.

`col.dis` *Object of class distance for Columbus dataset*

Description

Distance object, GWT and txt files for columbus dataset

Usage

```
coldis
```

Format

A list of neighbors and distances for the columbus neighborhoods distance *Writes distance matrices*

Description

Reads points coordinates and generates objects of class `distance.matrix`

Usage

```
distance(coord,region.id=NULL,output=TRUE,
type=c("NN","distance","inverse"),
measure=c("euclidean","gcircle","chebyshev","braycur","canberra"),
nn=6, cutoff=FALSE, miles=TRUE,R=NULL, shape.name=NULL,region.id.name=NULL,
firstline=FALSE,file.name=NULL)
```

Arguments

`coord`	a matrix with the (X,Y)-coordinates of the points. The first column can be the region.id variable giving the ordering of the observations
`region.id`	variable that defines the ordering of the observations
`output`	when TRUE (default) writes the object to a file
`type`	one of `("NN","distance","inverse")`. Nearest neighbors, distance or inverse distance
`measure`	one of `("euclidean","gcircle","chebyshev","braycur","canberra")`.The distance measure to be employed in the calculations.
`nn`	the number of nearest neighbors
`cutoff`	If type is `distance` or `inverse`. Assumes values 1, 2 or 3. When 1, the cutoff is set to the first quantile of the distribution of distances. When 2 to the median, and when 3 to the third quantile. Only observations with distance less than cutoff distance are neighbors.
`miles`	If TRUE (default), distances are in miles, otherwise in Km. (See `rdist.earth` for details)
`R`	Radius to use for sphere to find spherical distances.(See `rdist.earth` for details)
`shape.name`	The name of the shape file. See Details
`region.id.name`	The name of the `region.id` variable. See Details
`firstline`	If `TRUE`, a first line is added to the output file. See Details
`file.name`	If `output`, the name of the output file. See Details

Details

The object created is similar to the content of a `'GWT'` file. The output file can be of any format. In particular, it could be a `'GWT'` file. When `firstline` is `TRUE`, an header line is added to the `'GWT'` file. The first element is simply a place holder, the second is the number of observations. The name of the shape file and of the id variable can be specified by the options `shape.name` and `region.id.name` respectively. The function performs a series of test on the `region.id` variable. If a `region.id` variable is not specified and `coord` only has two columns, a sequence from 1 to the number of observations is generated and used as identification variable. If `region.id` is specified and the first column of `coord` contains an id variable they should be the same.

Value

A `matrix` of three columns: `from`, `to`, and `distance`

Author(s)

Gianfranco Piras <gpiras@mac.com>

Examples

```
X<-runif(100,0,70)
Y<-runif(100,-30,20)
coord1<-cbind(seq(1,100),X,Y)
thm1 <- distance(coord1,region.id=NULL,output=TRUE,type="inverse", shape.name="shapefile",region.id.name="id1",
thm2 <- distance(coord1,region.id=NULL,output=FALSE,type="NN", nn=6)
thm2 <- distance(coord1,region.id=NULL,output=FALSE,type="distance", cutoff=1)
```

`dist_functions` *Distance measures available in distance*

Description

The great circle distance is calculated by the function `rdist.earth` in library `fields`.

Details

The distance measures implemented in `sphet` are:

- `'euclidean'`: $\sqrt{\sum (x_i - y_i)^2}$
- `'chebyshev'`: $max(|x_i - y_i|)$
- `'braycur'`: $\frac{\sum |x_i - y_i|}{\sum |x_i + y_i|}$
- `'canberra'`: $\frac{\sum |x_i - y_i|}{\sum |x_i| + |y_i|}$
- `'gcircle'`: see `rdist.earth`.

Author(s)

Gianfranco Piras <gpiras@mac.com>

See Also

`gstslshet`, `distance`, `distance`

`gstslshet` *GM estimation of a Cliff-Ord type model with Heteroskedastic Innovations*

Description

Multi step GM/IV estimation of a linear Cliff and Ord -type of model of the form:

$$y = \lambda W y + X\beta + u$$

$$u = \rho W u + e$$

with

$$e \tilde{} N(0, \sigma_i^2)$$

The model allows for spatial lag in the dependent variable and disturbances. The innovations in the disturbance process are assumed heteroskedastic of an unknown form.

Usage

```
gstslshet(formula, data=list(),listw,na.action=na.fail,zero.policy=NULL,initial.value=0.2,abs.tol=1e
## S3 method for class 'gstsls'
impacts(obj, ..., tr, R = NULL, listw = NULL, tol = 1e-06, empirical = FALSE, Q=NULL)
```

Arguments

`formula`	a description of the model to be fit
`data`	an object of class data.frame. An optional data frame containing the variables in the model.
`listw`	an object of class `listw` created for example by `nb2listw`
`na.action`	a function which indicates what should happen when the data contains missing values. See lm for details.
`zero.policy`	See `lagsarlm` for details
`initial.value`	The initial value for ρ. It can be either numeric (default is 0.2) or set to `'SAR'`, in which case the optimization will start from the estimated coefficient of a regression of the 2SLS residuals over their spatial lag (i.e. a spatial AR model)
`abs.tol`	Absolute tolerance. See nlminb for details.
`rel.tol`	Relative tolerance. See nlminb for details.

`eps`	Tolerance level for the approximation. See Details.
`inverse`	`TRUE`. If `FALSE`, an appoximated inverse is calculated. See Details.
`sarar`	`TRUE`. If `FALSE`, a spatial error model is estimated.
`obj`	A gstsls spatial regression object created by `gstslshet`
`...`	Arguments passed through to methods in the **coda** package
`tr`	A vector of traces of powers of the spatial weights matrix created using `trW`, for approximate impact measures; if not given, `listw` must be given for exact measures (for small to moderate spatial weights matrices); the traces must be for the same spatial weights as were used in fitting the spatial regression
`R`	If given, simulations are used to compute distributions for the impact measures, returned as `mcmc` objects
`tol`	Argument passed to `mvrnorm`: tolerance (relative to largest variance) for numerical lack of positive-definiteness in the coefficient covariance matrix
`empirical`	Argument passed to `mvrnorm` (default FALSE): if true, the coefficients and their covariance matrix specify the empirical not population mean and covariance matrix
`Q`	default NULL, else an integer number of cumulative power series impacts to calculate if `tr` is given

Details

The procedure consists of two steps alternating GM and IV estimators. Each step consists of substeps. In step one $\delta = [\beta', \lambda]'$ is estimated by 2SLS. The 2SLS residuals are first employed to obtain an initial (consistent but not efficient) GM estimator of ρ and then a consistent and efficient estimator (involving the variance-covariance matrix of the limiting distribution of the normalized sample moments). In step two, the spatial Cochrane-Orcutt transformed model is estimated by 2SLS. This corresponds to a GS2SLS procedure. The GS2SLS residuals are used to obtain a consistent and efficient GM estimator for ρ.

The initial value for the optimization in step 1b is taken to be `initial.value`. The initial value in step 1c is the optimal parameter of step 1b. Finally, the initial value for the optimization of step 2b is the optimal parameter of step 1c.

Internally, the object of class `listw` is transformed into a Matrix using the function listw2dgCMatrix.

The expression of the estimated variance covariance matrix of the limiting distribution of the normalized sample moments based on 2SLS residuals involves the inversion of $I - \rho W'$. When `inverse` is `FALSE`, the inverse is calculated using the approximation $I + \rho W' + \rho^2 W'^2 + ... + \rho^n W'^n$. The powers considered depend on a condition. The function will keep adding terms until the absolute value of the `sum` of all elements of the matrix $\rho^i W^i$ is greater than a fixed ϵ (`eps`). By default `eps` is set to 1e-5.

Value

A list object of class `sphet`

`coefficients`	Generalized Spatial two stage least squares coefficient estimates of δ and GM estimator for ρ.

var	variance-covariance matrix of the estimated coefficients
s2	GS2SLS residuals variance
residuals	GS2SLS residuals
yhat	difference between GS2SLS residuals and response variable
call	the call used to create this object
model	the model matrix of data
method	'gs2slshac'
W	Wald test for both ρ and λ are zero

Author(s)

Gianfranco Piras <gpiras@mac.com>

References

Arraiz, I. and Drukker, M.D. and Kelejian, H.H. and Prucha, I.R. (2007) A spatial Cliff-Ord-type Model with Heteroskedastic Innovations: Small and Large Sample Results, *Department of Economics, University of Maryland*'

Kelejian, H.H. and Prucha, I.R. (2007) Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances, *Journal of Econometrics*, forthcoming.

Kelejian, H.H. and Prucha, I.R. (1999) A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model, *International Economic Review*, **40**, pages 509–533.

Kelejian, H.H. and Prucha, I.R. (1998) A Generalized Spatial Two Stage Least Square Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances, *Journal of Real Estate Finance and Economics*, **17**, pages 99–121.

See Also

stslshac

Examples

```
library(spdep)
data(columbus)
listw<-nb2listw(col.gal.nb)
res<-gstslshet(CRIME~HOVAL + INC, data=columbus, listw=listw)
summary(res)
```

kernel — *Spatial two stages least square with HAC standard errors*

Description

Main functions to calculate the kernels in `stslshac`

Details

Six different kernel functions are implemented:

- `'Epanechnikov'`: $K(z) = 1 - z^2$
- `'Triangular'`: $K(z) = 1 - z$
- `'Bisquare'`: $K(z) = (1 - z^2)^2$
- `'Parzen'`: $K(z) = 1 - 6z^2 + 6|z|^3$ if $z \leq 0.5$ and $K(z) = 2(1 - |z|)^3$ if $0.5 < z \leq 1$
- `'TH'` (Tukey - Hanning): $K(z) = \frac{1+\cos(\pi z)}{2}$
- `'QS'` (Quadratic Spectral): $K(z) = \frac{25}{12\pi^2 z^2}(\frac{\sin(6\pi z)/5)}{6\pi z/5} - \cos(6\pi z)/5))$.

If the kernel type is not one of the six implemented, the function will terminate with an error message. The spatial two stage least square estimator is based on the matrix of instruments $H = [X, WX, W^2X^2]$.

Author(s)

Gianfranco Piras <`gpiras@mac.com`>

See Also

`gstslshet`, `distance`, `distance`

listw2dgCMatrix — *Interface between Matrix class objects and weights list*

Description

Interface between Matrix class objects and weights list

Usage

```
listw2dgCMatrix(listw, zero.policy = NULL)
```

Arguments

`listw`	a `listw` object created for example by `nb2listw`
`zero.policy`	See `lagsarlm` for details

Value

Matrix class object: a sparse Matrix

Author(s)

Gianfranco Piras <gpiras@mac.com>

Examples

```
library(spdep)
data(columbus)
listw<-nb2listw(col.gal.nb)
spW<-listw2dgCMatrix(listw)
```

`print.sphet` *print method for class sphet*

Description

Method used to print objects of class 'summary.sphet' and 'sphet'

Usage

```
## S3 method for class 'sphet'
print(x, digits = max(3,getOption("digits") -3), ...)
```

Arguments

x	an object of class 'summary.sphet' and sphet
digits	minimal number of significant digits, see print.default
...	additional argument to be passed

Details

The summary function summary.sphet returns an objects of class 'sphet' organized in a coefficient matrix.

Author(s)

Gianfranco Piras<gpiras@mac.com>

See Also

gstslshet, stslshac

Examples

```
library(spdep)
data(columbus)
listw<-nb2listw(col.gal.nb)
res<-gstslshet(CRIME~HOVAL + INC, data=columbus, listw=listw)
summary(res)
```

`read.gwt2dist` *Read distance ojbects*

Description

The function reads "GWT" files (i.e. generated using distance. It will read also other more general formats (as for example .txt files).

Usage

```
read.gwt2dist(file, region.id=NULL, skip=1)
```

Arguments

`file`	name of file to be read
`region.id`	variable that defines the ordering of the observations
`skip`	skip number of lines

Details

The first line of a `'GWT'` file generally contains some information (e.g. the name of the shape file, the number of observations), in which case, `skip` should be equal to 1. When the `file` has a `'GWT'` extension, the number of observations is generally retrived from the first line. Alternatively, it is fixed to the length of the unique `region.id` variable.

Value

An object of class `distance`.

Author(s)

Gianfranco Piras <gpiras@mac.com>

Examples

```
## Not run: dist<-read.gwt2dist(file='knn10columbus.GWT', region.id=POLYID)
```

stslshac *Spatial two stages least square with HAC standard errors*

Description

Non-parametric heteroskedasticity and autocorrelation consistent (HAC) estimator of the variance-covariance (VC) for a vector of sample moments within a spatial context. The disturbance vector is generated as follows:

$$u = R\epsilon$$

where R is a non-stochastic matrix.

Usage

```
stslshac(formula, data=list(),listw,na.action=na.fail,zero.policy=NULL,
HAC=TRUE, distance=NULL,type=c("Epanechnikov","Triangular","Bisquare","Parzen", "QS","TH"),
bandwidth="variable",W2X=TRUE)
```

Arguments

`formula`	a description of the model to be fit
`data`	an object of class data.frame. An optional data frame containing the variables in the model.
`listw`	an object of class `listw` created for example by `nb2listw`
`distance`	an object of class `distance` created for example by read.gwt2dist The object contains the specification of the distance measure to be employed in the estimation of the VC matrix. See Details.
`type`	One of `c("Epanechnikov","Triangular","Bisquare","Parzen", "QS","TH")`. The type of Kernel to be used. See Details.
`na.action`	a function which indicates what should happen when the data contains missing values. See lm for details.
`zero.policy`	See `lagsarlm` for details
`bandwidth`	"variable" (default) - or numeric when a fixed bandwidth is specified by the user.
`HAC`	if FALSE traditional standard errors are provided.
`W2X`	default TRUE. if FALSE only WX are used as instruments in the spatial two stage least squares.

Details

The default sets the bandwith for each observation to the maximum distance for that observation (i.e. the max of each element of the list of distances).

Six different kernel functions are implemented:

- `'Epanechnikov'`: $K(z) = 1 - z^2$
- `'Triangular'`: $K(z) = 1 - z$

- 'Bisquare': $K(z) = (1 - z^2)^2$
- 'Parzen': $K(z) = 1 - 6z^2 + 6|z|^3$ if $z \leq 0.5$ and $K(z) = 2(1 - |z|)^3$ if $0.5 < z \leq 1$
- 'TH' (Tukey - Hanning): $K(z) = \frac{1+\cos(\pi z)}{2}$
- 'QS' (Quadratic Spectral): $K(z) = \frac{25}{12\pi^2 z^2}(\frac{\sin(6\pi z)/5)}{6\pi z/5} - \cos(6\pi z)/5))$.

If the kernel type is not one of the six implemented, the function will terminate with an error message. The spatial two stage least square estimator is based on the matrix of instruments $H = [X, WX, W^2X^2]$.

Value

A list object of class sphet

coefficients	Spatial two stage least squares coefficient estimates
vcmat	variance-covariance matrix of the estimated coefficients
s2	S2sls residulas variance
residuals	S2sls residuals
yhat	difference between residuals and response variable
call	the call used to create this object
model	the model matrix of data
type	the kernel employed in the estimation
bandwidth	the type of bandwidth
method	's2slshac'

Author(s)

Gianfranco Piras <gpiras@mac.com>

References

Kelejian, H.H. and Prucha, I.R. (2007) HAC estimation in a spatial framework, *Journal of Econometrics*, **140**, pages 131–154.

Kelejian, H.H. and Prucha, I.R. (1999) A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model, *International Economic Review*, **40**, pages 509–533.

Kelejian, H.H. and Prucha, I.R. (1998) A Generalized Spatial Two Stage Least Square Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances, *Journal of Real Estate Finance and Economics*, **17**, pages 99–121.

See Also

gstslshet, distance, distance

Examples

```
library(spdep)
data(columbus)
listw<-nb2listw(col.gal.nb)
data(coldis)
res<-stslshac(CRIME~HOVAL + INC, data=columbus,listw=listw, distance=coldis, type='Triangular')
summary(res)
```

`summary.sphet` *print method for class sphet*

Description

Method used to print objects of class 'summary.sphet' and 'sphet'

Usage

```
## S3 method for class 'sphet'
summary(object,width=getOption("width"), digits=getOption("digits"),obsinfo=FALSE,...)
```

Arguments

object	an object of class 'sphet'
width	controls the maximum number of columns on a line used in printing
digits	minimal number of significant digits, see print.default
obsinfo	for objects of class distance: if TRUE prints observation-wise information
...	additional arguments to be passed

Details

The summary function summary.sphet returns an objects of class 'sphet' organized in a coefficient matrix.

Author(s)

Gianfranco Piras<gpiras@mac.com>

See Also

gstslshet, stslshac

Examples

```
library(spdep)
data(columbus)
listw<-nb2listw(col.gal.nb)
res<-gstslshet(CRIME~HOVAL + INC, data=columbus, listw=listw)
summary(res)
```

utilities *Functions used by gstslshet.*

Description

- `arg` and `arg1` are the objective functions of the non-linear estimators in the GMM procedure.
- `Omega` and `Omegabis` generates the variance-covariance matrices of the Original and Transformed models (See Arraiz et al., 2007 for details.)
- `Ggfastfast` calculates G and g.
- All other functions perform calculations to estimates various objects defined in Appendix B2 and B3 in Arraiz et al., 2007.

Author(s)

Gianfranco Piras <gpiras@mac.com>

Index