Package 'eCAR'

May 15, 2021

Type Package

Title Eigenvalue CAR Models

Version 0.1.1

Maintainer Garritt L. Page <page@stat.byu.edu>

Description Fits Leroux model in spectral domain to estimate causal spatial effect as detailed in Guan, Y; Page, G.L.; Reich, B.J.; Ventrucci, M.; Yang, S; (2020) <arXiv:2012.11767>. Both the parametric and semi-parametric models are available. The semi-parametric model relies on 'INLA'. The 'INLA' package can be obtained from <https://www.r-inla.org/>.

License GPL

Encoding UTF-8

Depends R (>= 3.5.0)

Suggests INLA

Imports Matrix

LazyData true

RoxygenNote 7.1.1

Additional_repositories https://inla.r-inla-download.org/R/stable/

URL https://github.com/gpage2990/eCAR

NeedsCompilation yes

Author Garritt L. Page [aut, cre],
Massimo Ventrucci [aut],
S. McKay Curtis [cph],
Radford M. Neal [cph]

Repository CRAN

Date/Publication 2021-05-15 00:20:02 UTC

R topics documented:

`eCAR.out` *eCAR class constructor*

Description

A constructor for the `eCAR` class. The class `eCAR` is a named list containing the output from the calling the `par.eCAR.Leroux` or `semipar.eCAR.Leroux` functions.

Usage

```
eCAR.out(
  data_model = NULL,
  beta_omega = NULL,
  posterior_draws = NULL,
  DIC = NULL,
  regrcoef = NULL
)
```

Arguments

`data_model`	a characther indicating what data model was fit;
`beta_omega`	matrix containing estimated beta as a function of omega with 95% credible bands, and eigen-values;
`posterior_draws`	a list containing the posterior draws of all model parameters;
`DIC`	Deviance information criterion;
`regrcoef`	posterior summaries (mean, sd, 0.025quant, 0.975quant) for the regression coefs;

`lipcancer` *Number of recorded lip cancer cases in the 56 districts of Scotland.*

Description

A list containing the data frame and neighborhood matrix based on 56 districts of Scotland that is with the following variables

Usage

```
lipcancer
```

Format

data: A data frame with 56 rows and the following 6 variables:

observed observed number of cancer cases

expected the expected number of lip cancer cases computed using indirect standardisation based on Scotland-wide disease rates

pcaff percentage of the district's workforce employed in agriculture, fishing and forestry

latitude latitude coordinates

longitude longitude coordinates

name name

neighborhood.Matrix: A 56 x 56 matrix neighbhorhood matrix

`par.eCAR.Leroux` *R wrapper that accesses C code to fit parametric Leroux CAR spatial regression model that accounts for spatial confounding*

Description

`par.eCAR.Leroux` is the main function used to fit the parametric Leroux CAR model specified in the spectral domain.

Usage

```
par.eCAR.Leroux(y, x, W,
                E=NULL,
                C=NULL,
                model="Gaussian",
                joint_prior_lamx_lamz = FALSE,
                lamx.fix.val = NULL,
                sig2x.fix.val = NULL,
                m=0, s2=10,
                alamx=1, blamx=1,
                alamz=1, blamz=1,
                asig=1, bsig=1,
                atau=1, btau=1,
                asigx=1, bsigx=1,
                mb0=0,s2b0=100,
                me=0,s2e=100,
                mx=0,s2x=100,
                tau_cand_sd = 1,
                sig2_cand_sd = 1,
                draws=10000, burn=5000, thin=5,
                verbose=TRUE)
```

Arguments

y	response vector
x	covariate vector for which casual effect is desired
W	neighborhood matrix comprised of zeros and ones
E	This argument is ignored if model is Gaussian. For other models it takes on the following: • Poisson - E is vector that contains expected counts • Binomial - E is vector that contains number of trials • Negative Binomial - E is vector that contains an offset.
C	design matrix for the covariates that are included as controls
model	Specifies the likelihood or data model. Options are "Gaussian", "Poisson", "Binomial", "Negative Binomial"
joint_prior_lamx_lamz	Logical. If TRUE, then a uniform prior on space such that lamz > lamx. If FALSE, independent beta priors are used.
lamx.fix.val	If a value is supplied then lambda_x is not updated in the MCMC algorithm, but rather treated as the fixed known supplied value
sig2x.fix.val	If a value is supplied then sigma2_x is not updated in the MCMC algorithm, but rather treated as the fixed known supplied value
m	prior mean for (beta, gamma), where gamma = rho*(sigz/sigx). default is 0.
s2	prior variance for (beta, gamma),where gamma = rho*(sigz/sigx). default is 10
alamx	prior shape1 parameter for lam.x, default is 1. Only used if joint_prior_lamx_lamz = FALSE
blamx	prior shape2 parameter for lam.x, default is 1. Only used if joint_prior_lamx_lamz = FALSE
alamz	prior shape1 parameter for lam.z, default is 1. Only used if joint_prior_lamx_lamz = FALSE
blamz	prior shape2 parameter for lam.z, default is 1. Only used if joint_prior_lamx_lamz = FALSE
asig	prior shape parameter for sigma2, default is 1. Only used if model is Gaussian
bsig	prior scale parameter for sigma2, default is 1. Only used if model is Gaussian
atau	prior shape parameter for tau, where tau = sigma2.z*(1-rho^2). default is 1
btau	prior scale parameter for tau, where tau = sigma2.z*(1-rho^2). default is 1
asigx	prior shape parameter for sigma2.x, default is 1
bsigx	prior scale parameter for sigma2.x, default is 1
mb0	prior mean parameter for beta0, default is 0. Only used if model is not Gaussian
s2b0	prior variance parameter for beta0, default is 100. Only used if model is not Gaussian
me	prior mean parameter for eta, default is 0. Only used if C is not NULL
s2e	prior variance parameter for eta, default is 100. Only used if C is not NULL

mx	prior mean parameter for xi, default is 0. Only used for negative binomial model
s2x	prior variance parameter for eta, default is 100. Only used for negative binomial model
tau_cand_sd	standard deviation for candidate density in Metropolis step for tau. Default is 1
sig2_cand_sd	standard deviation for candidate density in Metropolis step for sig2. Default is 1. Only used if model is Gaussian
draws	number of MCMC iterates to be collected. default is 10000
burn	number of MCMC iterates discared as burn-in. default is 5000
thin	number by which the MCMC chain is thinned. default is 5
verbose	If TRUE, then details associated with data being fit are printed to screen along with MCMC iterate counter

Value

The function returns an eCAR object which is a list that contains the following

data_model	Character indicating which model was fit
beta_omega	Matrix that contains respectively, the posterior mean lower and upper quantiles of the (spatial scale)-varying beta at each omega value (for the non Gaussian cases it is the exponentiated beta).
posterior_draws	List containing posterior draws of the following parameters 1. beta: vector containing draws from the posterior distribution of exposure coefficient 2. gamma: vector containing draws from the posterior distribution of gamma=rho*(sigz/sigx) 3. tau: vector containing draws from tau=sigma2.z*(1-rho^2) 4. sig2x: vector containing draws from sig2x 5. lamx: vector containing draws from the posterior distribution of lamx 6. lamz: vector containing draws from the posterior distribution of lamz 7. sigma2: vector containing draws from the posterior distribution of sigma2. Only available if model is Gaussian 8. rho: vector containing draws from the posterior distribution of rho 9. sig2z: vector containing draws from the posterior distribution of sig2z 10. theta: matrix containing draws from the posterior distribution of theta. Only available if model is not Gaussian 11. beta0: vector containing draws from the posterior distribution of beta0. Only available if model is not Gaussian 12. eta: matrix containing draws from the posterior distribution of eta, regression coefficients of additional confouner variables. Only available if C is not NULL 13. nb_r: matrix containing draws from the posterior distribution of nb_r. 0nly available if model is Negative Binomial
DIC	Not available from parametric model yet
regrcoef	Not available from parametric model yet

References

Guan, Y; Page, G.L.; Reich, B.J.; Ventrucci, M.; Yang, S; "A spectral adjustment for spatial confounding" <arXiv:2012.11767>

Examples

```
# Our R-package
library(eCAR)

data(lipcancer)


W <- lipcancer$neighborhood.Matrix
M <- diag(apply(W,1,sum))
R <- M-W
e.dec <- eigen(R)
e.val <- e.dec$values
D.eigval = diag(e.val)


Y <- lipcancer$data$observed
X <- lipcancer$data$pcaff
E <- lipcancer$data$expected

set.seed(101)
fit1 <- par.eCAR.Leroux(y=Y, x=X, W=W, E=E, C=NULL, model="Poisson",
                        draws=10000, burn=5000, thin=1, verbose=FALSE,
                        joint_prior_lamx_lamz=FALSE, m=0, s2=4)


plot(fit1)
```

`plot` *Plot the spatial scale varying coefficient*

Description

This function takes the output obtained from the parametric or semiparametric fit and returns a plot of the spatial scale varying coefficient.

Usage

```
## S3 method for class 'eCAR'
plot(x, ...)
```

Arguments

x	an object of class eCAR (i.e. the output of the par.eCAR.Leroux() or semipar.eCAR.Leroux())
...	include here other inputs to the plot function

Value

This function returns the estimated posterior mean and 95-th credible intervals for the effect of the covariate of interest as a function of eigenvalues. If model is not Gaussian the exponential of the spatial scale varying coefficient is plotted which is useful in interpretating the covariate effect in the Binomial, Negative Binomial and Poisson models.

---

semipar.eCAR.Leroux *R wrapper that accesses 'INLA' to fit semi-parametric Leroux CAR spatial regression model that accounts for spatial confounding.*

Description

semipar.eCAR.Leroux is the main function used to fit the semi-parametric Leroux CAR model specified in the spectral domain. This function calls 'INLA'.

Usage

```
semipar.eCAR.Leroux(y, x, W,
                    E,
                    C=NULL,
                    names.covariates=NULL,
                    model="Gaussian",
                    L=10,
                    pcprior.sd=c(0.1,1),
                    s2=10,
                    method = "spectral",
                    num.threads.inla = NULL,
                    verbose=FALSE, ...)
```

Arguments

y	response vector
x	covariate vector for which casual effect is desired
W	neighborhood matrix comprised of zeros and ones
E	Offset value whose specification depends on the data model selected such that for * Poisson - E is vector that contains expected counts * Binomial - E is vector that contains number of trials * Negative Binomial - E is vector that contains an offset.

`C` design matrix for the covariates that are included as controls

`names.covariates`
Specifies the names of the covariates inside C

`model` Specifies the likelihood or data model. Options are "Gaussian", "Poisson", "Binomial", "Negative Binomial"

`L` Number of basis functions for the spline model on the (spatial scale)-varying beta. The smoothing method applied here is a Bayesian version of the P-spline approach by Eilers and Marx (1996), assuming a random walk on the spline coefficients and a PC-prior on the precision parameter of the random walk.

`pcprior.sd` Vector of length 2 specifying the scaling parameters for the PC-priors assumed on the precision of the (spatial scale)-varying beta and the data y, respectively. Each of the scaling parameters can be interpreted as a guess on the marginal standard deviation (default are 0.1 and 1).

`s2` Prior variance for the log of the dispersion parameter (only used for model="Negative Binomial", default equal to 10).

`method` A character defining the type of adjustment; either "spectral" (default choice) which implements the model assuming (spatial scale)-varying beta, or "naive" which implements the standard method with constant beta hence no spectral adjustment.

`verbose` logical; if TRUE the verbose output from the "inla" call is printed.

`num.threads.inla`
Argument that indicates the number of computing cores that the INLA call will occupy. For syntax, see "inla.setOption"

`...` Arguments to be passed to the "inla" call; for instance control.inla=list(strategy="laplace")

Value

A eCAR object which is a list containing the following

`data_model` Character indicating which model was fit

`beta_omega` Matrix that contains respectively, the posterior mean lower and upper quantiles of the (spatial scale)-varying beta at each omega value (for the non Gaussian cases it is the exponentiated beta).

`posterior_draws`
List containing posterior draws of the following parameters

1. postsample.beta: matrix of dim Lx1000, containing drawns from the posterior distribution of the L spline coefficients
2. postsample.prec.beta: vector of length 1000, containing drawns from the posterior distribution of the precision of the random walk on the spline coefficients
3. postsample.prec.z: vector of length 1000, containing drawns from the posterior distribution of prec.z
4. postsample.lambda.z: vector of length 1000, containing drawns from the posterior distribution of lambda.z

5. postsample.c.beta: vector of length 1000, containing drawns from the posterior distribution of the bias corrected beta (i.e. beta evaluated at the largest eigenvalue)

`DIC` Deviance information criterion computed by 'INLA'

`regrcoef` posterior summaries (mean, sd, 0.025quant, 0.975quant) for the regression coefs associated to the covariates inside 'C' (if model is Binomial, Poisson or Negative binomial, the posterior summaries refer to the exponentiated coef)

References

Guan, Y; Page, G.L.; Reich, B.J.; Ventrucci, M.; Yang, S; "A spectral adjustment for spatial confounding" <arXiv:2012.11767>

Examples

```
# Our R-package
library(eCAR)

data(lipcancer)
W <- lipcancer$neighborhood.Matrix
Y <- lipcancer$data$observed
X <- lipcancer$data$pcaff
E <- lipcancer$data$expected



if (requireNamespace("INLA", quietly = TRUE)) {
  fit1 = semipar.eCAR.Leroux(y=Y, x=X, W=W, E=E, C=NULL,
                                   pcprior.sd = c(0.1,1),
                                   model="Poisson",
                                   L=10,
                                   num.threads.inla = '1:1',
                                   verbose=FALSE)
  plot(fit1)
}
```

Index