Package: ref.ICAR (via r-universe)

September 15, 2024

Title Objective Bayes Intrinsic Conditional Autoregressive Model for Areal Data

Version 2.0.1

Author Erica M. Porter, Matthew J. Keefe, Christopher T. Franck, and Marco A.R. Ferreira

Maintainer Erica M. Porter <emporte@clemson.edu>

Description Implements an objective Bayes intrinsic conditional autoregressive prior. This model provides an objective Bayesian approach for modeling spatially correlated areal data using an intrinsic conditional autoregressive prior on a vector of spatial random effects.

License MIT + file LICENSE

Encoding UTF-8

Imports sf, sp, spdep, mvtnorm, coda, MCMCglmm, Rdpack, graphics, pracma, stats, classInt, dplyr, ggplot2, gtools

RdMacros Rdpack

Suggests maps, MASS, knitr, rmarkdown, RColorBrewer, rcrossref

VignetteBuilder knitr

BuildManual yes

RoxygenNote 7.2.3

NeedsCompilation no

Repository CRAN

Date/Publication 2023-08-22 08:50:02 UTC

Contents

`probs.icar` *OLM and ICAR model probabilities for areal data*

Description

Performs simultaneous selection of covariates and spatial model structure for areal data.

Usage

```
probs.icar(
  Y,
  X,
  H,
  H.spectral = NULL,
  Sig_phi = NULL,
  b = 0.05,
  verbose = FALSE
)
```

Arguments

`Y`	A vector of responses.
`X`	A matrix of covariates, which should include a column of 1's for models with a non-zero intercept
`H`	Neighborhood matrix for spatial subregions.
`H.spectral`	Spectral decomposition of neighborhood matrix, if user wants to pre-compute it to save time.
`Sig_phi`	Pseudo inverse of the neighborhood matrix, if user wants to pre-compute it to save time.
`b`	Training fraction for the fractional Bayes factor (FBF) approach.
`verbose`	If FALSE, marginal likelihood progress is not printed.

Value

A list containing a data frame with all posterior model probabilities and other selection information.

`probs.mat`	Data frame containing posterior model probabilities for all candidate OLMs and ICAR models from the data.
`mod.prior`	Vector of model priors used to obtain the posterior model probabilities.
`logmargin.all`	Vector of all (log) fractional integrated likelihoods.
`base.model`	Maximum (log) fractional integrated likelihood among all candidate models. All fractional Bayes factors are obtained with respect to this model.
`BF.vec`	Vector of fractional Bayes factors for all candidate models.

Author(s)

Erica M. Porter, Christopher T. Franck, and Marco A.R. Ferreira

References

Porter EM, Franck CT, Ferreira MAR (2023). "Objective Bayesian model selection for spatial hierarchical models with intrinsic conditional autoregressive priors." *Bayesian Analysis*, **1**(1), 1–27. doi:10.1214/23BA1375.

`ref.analysis` *MCMC Analysis and Summaries for Reference Prior on an Intrinsic Autoregressive Model for Areal Data*

Description

Performs analysis on a geographical areal data set using the objective prior for intrinsic conditional autoregressive (ICAR) random effects (Keefe et al. 2019). It takes a shapefile, data, and region names to construct a neighborhood matrix and perform Markov chain Monte Carlo sampling on the unstructured and spatial random effects. Finally, the function obtains regional estimates and performs posterior inference on the model parameters.

Usage

```
ref.analysis(
  shape.file,
  X,
  y,
  x.reg.names,
  y.reg.names,
  shp.reg.names = NULL,
  iters = 10000,
  burnin = 5000,
  verbose = TRUE,
  tauc.start = 1,
  beta.start = 1,
  sigma2.start = 1,
  step.tauc = 0.5,
  step.sigma2 = 0.5
)
```

Arguments

`shape.file`	A shapefile corresponding to the regions for analysis.
`X`	A matrix of covariates, which should include a column of 1's for models with a non-zero intercept
`y`	A vector of responses.

`x.reg.names`	A vector specifying the order of region names contained in `X`.
`y.reg.names`	A vector specifying the order of region names contained in `y`.
`shp.reg.names`	A vector specifying the order of region names contained in the shapefile, if there is not a NAME column in the file.
`iters`	Number of MCMC iterations to perform. Defaults to 10,000.
`burnin`	Number of MCMC iterations to discard as burn-in. Defaults to 5,000.
`verbose`	If FALSE, MCMC progress is not printed.
`tauc.start`	Starting MCMC value for the spatial dependence parameter.
`beta.start`	Starting MCMC value for the fixed effect regression coefficients.
`sigma2.start`	Starting MCMC value for the variance of the unstructured random effects.
`step.tauc`	Step size for the spatial dependence parameter.
`step.sigma2`	Step size for the variance of the unstructured random effects.

Value

A list containing `H`, MCMC chains, parameter summaries, fitted regional values, and regional summaries.

`H`	The neighborhood matrix.
`MCMC`	Matrix of MCMC chains for all model parameters.
`beta.median`	Posterior medians of the fixed effect regression coefficients.
`beta.hpd`	Highest Posterior Density intervals for the fixed effect regression coefficients.
`tauc.median`	Posterior median of the spatial dependence parameter.
`tauc.hpd`	Highest Posterior Density interval for the spatial dependence parameter.
`sigma2.median`	Posterior median of the unstructured random effects variance.
`sigma2.hpd`	Highest Posterior Density interval for the unstructured random effects variance.
`tauc.accept`	Final acceptance rate for the spatial dependence parameter.
`sigma2.accept`	Final acceptance rate for the unstructured random effects variance.
`fit.dist`	Matrix of fitted posterior values for each region in the data.
`reg.medians`	Vector of posterior medians for fitted response by region.
`reg.hpd`	Data frame of Highest Posterior Density intervals by region.

Author(s)

Erica M. Porter, Matthew J. Keefe, Christopher T. Franck, and Marco A.R. Ferreira

Examples

```
## Refer to the vignette attached to the package.
```

ref.MCMC *MCMC for Reference Prior on an Intrinsic Conditional Autoregressive Random Effects Model for Areal Data*

Description

Implements the Metropolis-within-Gibbs sampling algorithm proposed by Ferreira et al. (2021), to perform posterior inference for the intrinsic conditional autoregressive model with spatial random effects. This algorithm uses the spectral domain for the hierarchical model to create the Spectral Gibbs Sampler (SGS), which provides notable speedups to the MCMC algorithm proposed by Keefe et al (2019).

Usage

```
ref.MCMC(
  y,
  X,
  H,
  iters = 10000,
  burnin = 5000,
  verbose = TRUE,
  tauc.start = 1,
  beta.start = 1,
  sigma2.start = 1,
  step.tauc = 0.5,
  step.sigma2 = 0.5
)
```

Arguments

`y`	Vector of responses.
`X`	Matrix of covariates. This should include a column of 1's for models with a non-zero intercept.
`H`	The neighborhood matrix.
`iters`	Number of MCMC iterations to perform. Defaults to 10,000.
`burnin`	Number of MCMC iterations to discard as burn-in. Defaults to 5,000.
`verbose`	If FALSE, MCMC progress is not printed.
`tauc.start`	Starting value for the spatial dependence parameter.
`beta.start`	Starting value for the vector of fixed effect regression coefficients.
`sigma2.start`	Starting value for the variance of the unstructured random effects.
`step.tauc`	Step size for the spatial dependence parameter
`step.sigma2`	Step size for the variance of the unstructured random effects.

Value

A list containing MCMC chains and parameter summaries.

`MCMCchain`	Matrix of MCMC chains.
`tauc.MCMC`	MCMC chains for the spatial dependence parameter.
`sigma2.MCMC`	MCMC chains for the variance of the unstructured random effects.
`phi.MCMC`	MCMC chains for the spatial random effects.
`beta.MCMC`	MCMC chains for the fixed effect regression coefficients.
`accept.sigma2`	Final acceptance number for variance of the unstructured random effects.
`accept.tauc`	Final acceptance number for spatial dependence parameter.
`accept.phi`	Final acceptance number for spatial random effects.

Author(s)

Erica M. Porter, Matthew J. Keefe, Christopher T. Franck, and Marco A.R. Ferreira

References

Keefe MJ, Ferreira MAR, Franck CT (2019). "Objective Bayesian analysis for Gaussian hierarchical models with intrinsic conditional autoregressive priors." *Bayesian Analysis*, **14**(1), 181 – 209. doi:10.1214/18BA1107.

Keefe MJ, Ferreira MAR, Franck CT (2018). "On the formal specification of sum-zero constrained intrinsic conditional autoregressive models." *Spatial Statistics*, **24**, 54–65. doi:10.1016/j.spasta.2018.03.007.

Ferreira MAR, Porter EM, Franck CT (2021). "Fast and scalable computations for Gaussian hierarchical models with intrinsic conditional autoregressive spatial random effects." *Computational Statistics and Data Analysis*, **162**, 107264. ISSN 0167-9473, doi:10.1016/j.csda.2021.107264, `https://www.sciencedirect.com/science/article/pii/S0167947321000980`.

Examples

```
#### Fit the model for simulated areal data on a grid ####

### Load extra libraries
library(sp)
library(methods)
library(spdep)
library(mvtnorm)

### Generate areal data on a grid
rows=5; cols=5
tauc=1
sigma2=2; beta=c(1,5)

### Create grid
grid <- GridTopology(c(1,1), c(1,1), c(cols,rows))
polys <- as(grid, "SpatialPolygons")
```

```
spgrid <- SpatialPolygonsDataFrame(polys,data=data.frame(row.names=row.names(polys)))

### Create neighborhood matrix
grid.nb <- poly2nb(spgrid,queen=FALSE)
W <- nb2mat(grid.nb, style="B")

### Put spatially correlated data in grid
p <- length(beta)
num.reg <- (rows*cols)
if(p>1){x1<-rmvnorm(n=num.reg,mean=rep(0,p-1),sigma=diag(p-1))} else{x1<-NULL}
X <- cbind(rep(1,num.reg),x1)
Dmat <- diag(apply(W,1,sum))
H <- Dmat - W
row.names(H) <- NULL

### Obtain true response vector
theta_true <- rnorm(num.reg,mean=0,sd=sqrt(sigma2))
Q <- eigen(H,symmetric=TRUE)$vectors
eigH <- eigen(H,symmetric=TRUE)$values
D <- diag(eigH)
Qmat <- Q[,1:(num.reg-1)]
phimat <- diag(1/sqrt(eigH[1:(num.reg-1)]))
z <- t(rmvnorm(1,mean=rep(0,num.reg-1),sigma=diag(num.reg-1)))
phi_true <- sqrt((1/tauc)*sigma2)*(Qmat%*%phimat%*%z)
Y <- X%*%beta + theta_true + phi_true

### Fit the model
set.seed(5432)
model <- ref.MCMC(y=Y,X=X,H=H,iters=15000,burnin=5000,verbose=TRUE,tauc.start=.1,beta.start=-1,
sigma2.start=.1,step.tauc=0.5,
      step.sigma2=0.5)


#### Small example for checking
model <- ref.MCMC(y=Y,X=X,H=H,iters=1000,burnin=50,verbose=TRUE,tauc.start=.1,beta.start=-1,
sigma2.start=.1,step.tauc=0.5,
      step.sigma2=0.5)
```

---

ref.plot *Trace Plots for Parameters in ICAR Model*

Description

This function creates trace plots for the parameters in the ICAR reference prior model (Keefe et al. 2019).

Usage

```
ref.plot(MCMCchain, X, burnin, num.reg)
```

Arguments

`MCMCchain`	Matrix of MCMC chains for the model parameters.
`X`	Matrix of covariates.
`burnin`	Number of MCMC iterations from `MCMCchain` discarded as burn-in.
`num.reg`	Number of regions in the areal data set.

Value

Trace plots for the fixed effect regression coefficients, the precision parameter, and the unstructured random effects variance.

Author(s)

Erica M. Porter, Matthew J. Keefe, Christopher T. Franck, and Marco A.R. Ferreira

Examples

```
## Refer to the vignette attached to the package.
```

`ref.summary` *Parameter Summaries for MCMC Analysis*

Description

Takes a matrix of MCMC chains, iterations, and acceptance values to return posterior summaries of the parameters, including posterior medians, intervals, and acceptance rates.

Usage

```
ref.summary(
  MCMCchain,
  tauc.MCMC,
  sigma2.MCMC,
  beta.MCMC,
  phi.MCMC,
  accept.phi,
  accept.sigma2,
  accept.tauc,
  iters = 10000,
  burnin = 5000
)
```

Arguments

`MCMCchain`	Matrix of MCMC chains for the ICAR model parameters.
`tauc.MCMC`	MCMC chains for the spatial dependence parameter.
`sigma2.MCMC`	MCMC chains for the variance of the unstructured random effects.
`beta.MCMC`	MCMC chains for the fixed effect regression coefficients.
`phi.MCMC`	MCMC chains for the spatial random effects.
`accept.phi`	Final acceptance number for spatial random effects.
`accept.sigma2`	Final acceptance number for variance of the unstructured random effects.
`accept.tauc`	Final acceptance number for the spatial dependence parameter.
`iters`	Number of MCMC iterations in `MCMCchain`.
`burnin`	Number of MCMC iterations discarded as burn-in for `MCMCchain`.

Value

Parameter summaries

`beta.median`	Posterior medians of the fixed effect regression coefficients.
`beta.hpd`	Highest Posterior Density intervals for the fixed effect regression coefficients.
`tauc.median`	Posterior median of the spatial dependence parameter.
`tauc.hpd`	Highest Posterior Density interval for the spatial dependence parameter.
`sigma2.median`	Posterior median of the unstructured random effects variance.
`sigma2.hpd`	Highest Posterior Density interval for the unstructured random effects variance.
`tauc.accept`	Final acceptance rate for the spatial dependence parameter.
`sigma2.accept`	Final acceptance rate for the unstructured random effects variance.

Author(s)

Erica M. Porter, Matthew J. Keefe, Christopher T. Franck, and Marco A.R. Ferreira

Examples

```
## Refer to the vignette attached to the package.
```

`reg.summary` *Regional Summaries for Areal Data Modeled by ICAR Reference Prior Model*

Description

This function takes data and sampled MCMC chains for an areal data set and gives fitted posterior values and summaries by region using the model by (Keefe et al. 2019).

Usage

```
reg.summary(MCMCchain, X, Y, burnin)
```

Arguments

`MCMCchain`	Matrix of MCMC chains, using the sampling from (Keefe et al. 2019).
`X`	Matrix of covariates.
`Y`	Vector of responses.
`burnin`	Number of MCMC iterations discarded as burn-in in `MCMCchain`.

Value

A list of the fitted distributions by region, and medians and credible intervals by region.

`fit.dist`	Matrix of fitted posterior values for each region in the data.
`reg.medians`	Vector of posterior medians for fitted response by region.
`reg.cred`	Data frame of credbile intervals by region.

Author(s)

Erica M. Porter, Matthew J. Keefe, Christopher T. Franck, and Marco A.R. Ferreira

Examples

```
## Refer to the vignette attached to the package.
```

shape.H *Creating a Neighborhood Matrix for Areal Data from a Shapefile*

Description

Takes a path to a shape file and creates a neighborhood matrix. This neighborhood matrix can be used with the objective ICAR model (Keefe et al. 2018).

Usage

```
shape.H(shape.file)
```

Arguments

`shape.file`	File path to a shapefile.

Value

A list containing a neighborhood matrix and the SpatialPolygonsDataFrame object corresponding to the shape file.

`H`	A neighborhood matrix.
`map`	SpatialPolygonsDataFrame object from the provided shapefile.

Author(s)

Erica M. Porter, Matthew J. Keefe, Christopher T. Franck, and Marco A.R. Ferreira

Examples

```
#### Load extra libraries
library(sp)
library(sf)

### Read in a shapefile of the contiguous U.S. from package data
system.path <- system.file("extdata", "us.shape48.shp", package = "ref.ICAR", mustWork = TRUE)
shp.layer <- gsub('.shp','',basename(system.path))
shp.path <- dirname(system.path)
us.shape48 <- st_read(dsn = path.expand(shp.path), layer = shp.layer)

shp.data <- shape.H(system.path)
names(shp.data)
```

Index