Package ‘clespr’

October 12, 2022

Type Package

Title Composite Likelihood Estimation for Spatial Data

Version 1.1.2

Maintainer Ting Fung (Ralph) Ma <tingfung.ma@wisc.edu>

Description Composite likelihood approach is implemented to estimating statistical models for spatial ordinal and proportional data based on Feng et al. (2014) <doi:10.1002/env.2306>. Parameter estimates are identified by maximizing composite log-likelihood functions using the limited memory BFGS optimization algorithm with bounding constraints, while standard errors are obtained by estimating the Godambe information matrix.

License GPL-2

LazyData TRUE

RoxygenNote 6.0.1

Depends R (>= 3.2.0)

Imports AER (>= 1.2-5), pbivnorm (>= 0.6.0), MASS (>= 7.3-45), magic (>= 1.5-6), survival (>= 2.37-5), clordr (>= 1.0.2), doParallel (>= 1.0.11), foreach (>= 1.2.0), utils, stats

NeedsCompilation no

Author Ting Fung (Ralph) Ma [cre, aut],
Wenbo Wu [aut],
Jun Zhu [aut],
Xiaoping Feng [aut],
Daniel Walsh [ctb],
Robin Russell [ctb]

Repository CRAN

Date/Publication 2018-02-23 19:05:06 UTC

R topics documented:

`func.cl.ord` *Composite Likelihood Calculation for Spatial Ordinal Data*

Description

`func.cl.ord` calculates the composite log-likelihood for spatial ordered probit models.

Usage

```
func.cl.ord(vec.yobs, mat.X, mat.lattice, radius, n.cat, vec.par)
```

Arguments

`vec.yobs`	a vector of observed responses for all N sites.
`mat.X`	regression (design) matrix, including intercepts.
`mat.lattice`	a data matrix containing geographical information of sites. The *i* th row constitutes a set of geographical coordinates.
`radius`	weight radius.
`n.cat`	number of categories, at least 2.
`vec.par`	a vector of parameters consecutively as follows: a series of cutoffs (excluding -Inf, 0 and Inf) for latent responses, a vector of covariate parameters, a parameter 'sigmasq' modeling covariance matrix, 0<=sigmasq<=1, and a parameter 'rho' reflecting spatial correlation, abs(rho)<=1.

Value

`func.cl.ord` returns a list: number of categories, sum of weights, composite log-likelihood, a vector of scores, and a matrix of first-order partial derivatives for `vec.par`.

References

Feng, Xiaoping, Zhu, Jun, Lin, Pei-Sheng, and Steen-Adams, Michelle M. (2014) Composite likelihood Estimation for Models of Spatial Ordinal Data and Spatial Proportional Data with Zero/One values. *Environmetrics* 25(8): 571–583.

Examples

```
# True parameter
vec.cutoff <- 2; vec.beta <- c(1, 2, 1, 0, -1); sigmasq <- 0.8; rho <- 0.6; radius <- 5
vec.par <- c(vec.cutoff, vec.beta, sigmasq, rho)

# Coordinate matrix
n.cat <- 3; n.lati <- 30; n.long <- 30
n.site <- n.lati * n.long
mat.lattice <- cbind(rep(1:n.lati, n.long), rep(1:n.long, each=n.lati))
mat.dist <- as.matrix(dist(mat.lattice, upper=TRUE, diag=TRUE))
mat.cov <- sigmasq * rho^mat.dist

set.seed(1228)
# Generate regression (design) matrix with intercept
mat.X <- cbind(rep(1, n.site),scale(matrix(rnorm(n.site*(length(vec.beta)-1)),nrow=n.site)))
vec.Z <- t(chol(mat.cov)) %*% rnorm(n.site) + mat.X %*% vec.beta
vec.epsilon <- diag(sqrt(1-sigmasq), n.site) %*% rnorm(n.site)
vec.ylat <- as.numeric(vec.Z + vec.epsilon)

# Convert to the vector of observation
vec.yobs <- func.obs.ord(vec.ylat, vec.alpha=c(-Inf,0,vec.cutoff,Inf))

# Using func.cl.ord()
ls <- func.cl.ord(vec.yobs, mat.X, mat.lattice, radius, n.cat, vec.par)
ls$log.lkd
```

---

func.cl.ord.repar — *Reparameterized Composite Likelihood Calculation for Spatial Ordinal Data*

Description

`func.cl.ord` calculates the composite log-likelihood for reparameterized spatial ordered probit models. This function is internally called by `func.cle.ord`.

Usage

```
func.cl.ord.repar(vec.yobs, mat.X, mat.lattice, radius, n.cat, vec.repar)
```

Arguments

`vec.yobs`	a vector of observed responses for all N sites.
`mat.X`	regression (design) matrix, including intercepts.
`mat.lattice`	a data matrix containing geographical information of sites. The *i* th row constitutes a set of geographical coordinates.
`radius`	weight radius.

n.cat	number of categories, at least 2.
vec.repar	a vector of parameters consecutively as follows: a reparameterized vector (tau's) for latent responses, a vector of covariate parameters, a parameter 'sigmasq' modeling covariance matrix, 0<=sigmasq<=1, and a parameter 'rho' reflecting spatial correlation, abs(rho)<=1.

Value

func.cl.ord returns a list: number of categories, sum of weights, composite log-likelihood, a vector of scores, and a matrix of first-order partial derivatives for vec.par.

References

Feng, Xiaoping, Zhu, Jun, Lin, Pei-Sheng, and Steen-Adams, Michelle M. (2014) Composite likelihood Estimation for Models of Spatial Ordinal Data and Spatial Proportional Data with Zero/One values. *Environmetrics* 25(8): 571–583.

---

func.cl.prop *Composite Likelihood Calculation for Spatial Proportional Data*

Description

func.cl.prop calculates the composite log-likelihood for spatial Tobit models.

Usage

```
func.cl.prop(vec.yobs, mat.X, mat.lattice, radius, vec.par)
```

Arguments

vec.yobs	a vector of observed responses for all N sites.
mat.X	regression (design) matrix, including intercepts.
mat.lattice	a data matrix containing geographical information of sites. The *i* th row constitutes a set of geographical coordinates.
radius	weight radius.
vec.par	a vector of parameters consecutively as follows: a cutoff point for latent responses, a vector of covariate parameters, a parameter 'sigmasq' modeling covariance matrix, 0<=sigmasq<=1, and a parameter 'rho' reflecting spatial correlation, abs(rho)<=1.

Value

func.cl.prop returns a list of sum of weights, composite log-likelihood, a vector of scores, and a matrix of first-order partial derivatives for vec.par.

References

Feng, Xiaoping, Zhu, Jun, Lin, Pei-Sheng, and Steen-Adams, Michelle M. (2014) Composite likelihood Estimation for Models of Spatial Ordinal Data and Spatial Proportional Data with Zero/One values. *Environmetrics* 25(8): 571–583.

Examples

```
# True parameter
alpha <- 4; vec.beta <- c(1, 2, 1, 0, -1); sigmasq <- 0.8; rho <- 0.6; radius <- 5
vec.par <- c(alpha, vec.beta, sigmasq, rho)

# Coordinate matrix
n.lati <- 30; n.long <- 30
n.site <- n.lati * n.long
mat.lattice <- cbind(rep(1:n.lati, n.long), rep(1:n.long, each=n.lati))
mat.dist <- as.matrix(dist(mat.lattice, upper=TRUE, diag=TRUE))
mat.cov <- sigmasq * rho^mat.dist

set.seed(1228)

# Generate regression (design) matrix with intercept
mat.X <- cbind(rep(1, n.site),scale(matrix(rnorm(n.site*(length(vec.beta)-1)),nrow=n.site)))
vec.Z <- t(chol(mat.cov)) %*% rnorm(n.site) + mat.X %*% vec.beta
vec.epsilon <- diag(sqrt(1-sigmasq), n.site) %*% rnorm(n.site)
vec.ylat <- as.numeric(vec.Z + vec.epsilon)

# Convert to the vector of observation
vec.yobs <- func.obs.prop(vec.ylat, alpha=alpha)

# Use func.cl.prop()
ls <- func.cl.prop(vec.yobs, mat.X, mat.lattice, radius, vec.par)
ls$log.lkd
```

`func.cle.ord` *Composite Likelihood Estimation for Spatial Ordinal Data*

Description

`func.cle.ord` performs composite likelihood estimation of parameters and their standard errors in a spatial ordered probit model by maximizing its composite log-likelihood.

Usage

```
func.cle.ord(vec.yobs, mat.X, mat.lattice, radius, n.cat, n.sim = 100,
  parallel = TRUE, n.core = max(detectCores()/2, 1), output = TRUE)
```

Arguments

- `vec.yobs` a vector of observed responses for all N sites.
- `mat.X` regression (design) matrix, including intercepts.
- `mat.lattice` a data matrix containing geographical information of sites. The ith row constitutes a set of geographical coordinates.
- `radius` weight radius.
- `n.cat` number of categories.
- `n.sim` number of simulations used for calculate the Godambe matrix (default: 100).
- `parallel` logical flag indicates using parallel processing (default: `TRUE`).
- `n.core` number of physical cores used for parallel processing (when `parallel` is `TRUE`, default value is `max(detectCores()/2,1)`).
- `output` logical flag indicates whether printing out result (default: `TRUE`).

Details

Given the design matrix, the vector of observed responses, spatial lattice data, weight radius, number of categories, and the prespecified number of simulated vectors of responses used in estimating the Godambe information, this function assumes initial values of cutoff points and β as the estimates from the standard ordered probit regression with independent responses. After initial reparameterization, it first estimates parameters of interest by maximizing the composite log-likelihood using `optim`, then computes the reparameterized sample covariance matrix and the set of standard errors, and finally reverse the reparameterization to obtain estimates corresponding to the original parameterization.

Value

`func.cle.ord` returns a list containing:

`vec.par`: a vector of estimator for θ=(cutoff,β, σ^2, ρ);

`vec.se`: a vector of standard error for the estimator;

`mat.asyvar`: estimated asymptotic covariance matrix $H^{-1}(\theta)J(\theta)H^{-1}(\theta)$ for the estimator; and

`vec.comp`: a vector of computational time for parameter and standard error estimation.

`CLIC`: Composite likelihood information criterion proposed by Varin and Vidoni (2005), i.e. $-2 * logCL(\theta) + 2 * trace(H^{-1}(\theta)J(\theta))$

References

Feng, Xiaoping, Zhu, Jun, Lin, Pei-Sheng, and Steen-Adams, Michelle M. (2014) Composite likelihood Estimation for Models of Spatial Ordinal Data and Spatial Proportional Data with Zero/One values. *Environmetrics* 25(8): 571–583.

Examples

```
# Example of n.cat = 3 (Spatial ordinal regression)
# True parameter
vec.cutoff <- 2; vec.beta <- c(1, 2, 1, 0, -1); sigmasq <- 0.8; rho <- 0.6; radius <- 5
vec.par <- c(vec.cutoff, vec.beta, sigmasq, rho)

# Coordinate matrix
n.cat <- 3; n.lati <- 30; n.long <- 30
n.site <- n.lati * n.long
mat.lattice <- cbind(rep(1:n.lati, n.long), rep(1:n.long, each=n.lati))
mat.dist <- as.matrix(dist(mat.lattice, upper=TRUE, diag=TRUE))
mat.cov <- sigmasq * rho^mat.dist

set.seed(1228)
# Generate regression (design) matrix with intercept
mat.X <- cbind(rep(1, n.site),scale(matrix(rnorm(n.site*(length(vec.beta)-1)),nrow=n.site)))
vec.Z <- t(chol(mat.cov)) %*% rnorm(n.site) + mat.X %*% vec.beta
vec.epsilon <- diag(sqrt(1-sigmasq), n.site) %*% rnorm(n.site)
vec.ylat <- as.numeric(vec.Z + vec.epsilon)

# Convert to the vector of observation
vec.yobs <- func.obs.ord(vec.ylat, vec.alpha=c(-Inf,0,vec.cutoff,Inf))

# With parallel computing

## Not run:
ord.example <- func.cle.ord(vec.yobs, mat.X, mat.lattice, radius, n.cat,
n.sim=100, parallel = TRUE, n.core = 2)

round(ord.example$vec.par,4)
# alpha1   beta0   beta1   beta2   beta3   beta4 sigma^2     rho
# 1.8395  0.9550  1.9690  0.9565  0.0349 -1.0398  0.8200  0.5578

round(ord.example$vec.se,4)
# alpha1   beta0   beta1   beta2   beta3   beta4 sigma^2     rho
# 0.1602  0.1222  0.1463  0.0916  0.0485  0.0889  0.1935  0.1267

## End(Not run)

# Without parallel computing

## Not run:
ord.example2 <- func.cle.ord(vec.yobs, mat.X, mat.lattice, radius,
n.cat, n.sim=100, parallel = FALSE)

## End(Not run)

# Example for n.cat = 2 (i.e. Spatial probit regression)
# True parameter
vec.beta <- c(1, 2, 1, 0, -1); sigmasq <- 0.5; rho <- 0.6; radius <- 5
vec.par <- c(vec.beta, sigmasq, rho)
```

```
# Coordinate matrix
n.cat <- 2 ; n.lati <- n.long <- 40
n.site <- n.lati * n.long
mat.lattice <- cbind(rep(1:n.lati, n.long), rep(1:n.long, each=n.lati))
mat.dist <- as.matrix(dist(mat.lattice, upper=TRUE, diag=TRUE))
mat.cov <- sigmasq * rho^mat.dist

set.seed(123)
# Generate regression (design) matrix with intercept
mat.X <- cbind(rep(1, n.site),scale(matrix(rnorm(n.site*(length(vec.beta)-1)),nrow=n.site)))
vec.Z <- t(chol(mat.cov)) %*% rnorm(n.site) + mat.X %*% vec.beta
vec.epsilon <- diag(sqrt(1-sigmasq), n.site) %*% rnorm(n.site)
vec.ylat <- as.numeric(vec.Z + vec.epsilon)
# Convert to the vector of observation
vec.yobs <- func.obs.ord(vec.ylat, vec.alpha=c(-Inf,0,Inf))

## Not run:
probit.example <- func.cle.ord(vec.yobs, mat.X, mat.lattice, radius, n.cat,
n.sim=100, parallel = TRUE, n.core = 4)

round(probit.example$vec.par,4)
# beta0   beta1   beta2   beta3   beta4 sigma^2     rho
# 1.0427  2.2250  1.0422  0.0156 -1.1489  0.4402  0.6636

round(probit.example$vec.se,4)
# beta0   beta1   beta2   beta3   beta4 sigma^2     rho
# 0.1198  0.1413  0.0863  0.0523  0.0935  0.1600  0.1263

## End(Not run)
```

`func.cle.prop` *Composite Likelihood Estimation for Spatial Proportional Data*

Description

`func.cle.prop` performs composite likelihood estimation of parameters and their standard errors in a spatial Tobit model by maximizing its composite log-likelihood.

Usage

```
func.cle.prop(vec.yobs, mat.X, mat.lattice, radius, n.sim = 100,
  parallel = TRUE, n.core = max(detectCores()/2, 1), output = TRUE)
```

Arguments

`vec.yobs`	a vector of observed responses for all N sites.
`mat.X`	regression (design) matrix, including intercepts.

`mat.lattice`	a data matrix containing geographical information of sites. The i-th row constitutes a set of geographical coordinates.
`radius`	weight radius.
`n.sim`	number of simulations used for calculate the Godambe matrix (default: 100).
`parallel`	logical flag indicating using parallel processing (default: `TRUE`)
`n.core`	number of physical cores used for parallel processing (when `parallel` is `TRUE`), default value is `max(detectCores()/2,1)`).
`output`	logical flag indicates whether printing out result (default: `TRUE`).

Details

Given the design matrix, the vector of observed responses, spatial lattice data, weight radius, and the prespecified number of simulated vectors of responses used in estimating the Godambe information matrix, this function assumes initial values of β as the estimates from the standard Type I Tobit model with independent responses. The initial value of α and the right limit of the Tobit model are equally set to 1. Since there is only one cutoff point to be estimated, reparameterization is unnecessary. The function first estimates parameters of interest by maximizing the composite log-likelihood using `optim(...,method = "L-BFGS-B")`, then computes the simulated based standard error and asymptotic covariance matrix.

Value

`func.cle.prop` returns a list containing:

`vec.par`: a vector of estimator for $\theta = (\alpha, \beta, \sigma^2, \rho)$;

`vec.se`: a vector of standard error for the estimator;

`mat.asyvar`: estimated asymptotic covariance matrix $H^{-1}(\theta)J(\theta)H^{-1}(\theta)$ for the estimator; and

`vec.comp`: a vector of computational time for parameter and standard error estimation.

`CLIC`: Composite likelihood information criterion proposed by Varin and Vidoni (2005), i.e. $-2 * logCL(\theta) + 2 * trace(H^{-1}(\theta)J(\theta))$

References

Feng, Xiaoping, Zhu, Jun, Lin, Pei-Sheng, and Steen-Adams, Michelle M. (2014) Composite likelihood Estimation for Models of Spatial Ordinal Data and Spatial Proportional Data with Zero/One values. *Environmetrics* 25(8): 571–583.

Examples

```
# True parameter
alpha <- 4; vec.beta <- c(1, 2, 1, 0, -1); sigmasq <- 0.8; rho <- 0.6; radius <- 5
vec.par <- c(alpha, vec.beta, sigmasq, rho)

# Coordinate matrix
n.lati <- 30; n.long <- 30
n.site <- n.lati * n.long
mat.lattice <- cbind(rep(1:n.lati, n.long), rep(1:n.long, each=n.lati))
mat.dist <- as.matrix(dist(mat.lattice, upper=TRUE, diag=TRUE))
```

```
mat.cov <- sigmasq * rho^mat.dist

set.seed(1228)

# Generate regression (design) matrix with intercept
mat.X <- cbind(rep(1, n.site),scale(matrix(rnorm(n.site*(length(vec.beta)-1)),nrow=n.site)))
vec.Z <- t(chol(mat.cov)) %*% rnorm(n.site) + mat.X %*% vec.beta
vec.epsilon <- diag(sqrt(1-sigmasq), n.site) %*% rnorm(n.site)
vec.ylat <- as.numeric(vec.Z + vec.epsilon)

# Convert to the vector of observation
vec.yobs <- func.obs.prop(vec.ylat, alpha=alpha)

# With parallel computing

## Not run:
prop.example <- func.cle.prop(vec.yobs, mat.X, mat.lattice, radius,
n.sim=100, parallel = TRUE, n.core = 2)

round(prop.example$vec.par,4)
# alpha   beta0   beta1   beta2   beta3   beta4 sigma^2     rho
# 3.8259  0.9921  1.9679  0.9455  0.0148 -0.9871  0.8386  0.5761

round(prop.example$vec.se ,4)
# alpha   beta0   beta1   beta2   beta3   beta4 sigma^2     rho
# 0.1902  0.1406  0.1103  0.0744  0.0385  0.0652  0.1527  0.1151

## End(Not run)

# Without parallel computing

## Not run:
prop.example2 <- func.cle.prop(vec.yobs, mat.X, mat.lattice, radius, n.sim=100, parallel = FALSE)

## End(Not run)
```

`func.obs.ord` *Latent Response Transformation for Spatial Ordinal Data*

Description

`func.obs.ord` transforms a vector of latent responses into the corresponding observed ones under the spatial Probit model.

Usage

```
func.obs.ord(vec.ylat, vec.alpha)
```

Arguments

vec.ylat a vector of latent responses for all N sites.

vec.alpha a vector of prespecified cutoff points, ascending with length at least 3, including -Inf, 0, and Inf.

Value

func.obs.prop returns a vector of observed responses.

References

Feng, Xiaoping, Zhu, Jun, Lin, Pei-Sheng, and Steen-Adams, Michelle M. (2014) Composite likelihood Estimation for Models of Spatial Ordinal Data and Spatial Proportional Data with Zero/One values. *Environmetrics* 25(8): 571–583.

Examples

```
# True parameter
vec.cutoff <- 2; vec.beta <- c(1, 2, 1, 0, -1); sigmasq <- 0.8; rho <- 0.6; radius <- 5
vec.par <- c(vec.cutoff, vec.beta, sigmasq, rho)

# Coordinate matrix
n.cat <- 3; n.lati <- 30; n.long <- 30
n.site <- n.lati * n.long
mat.lattice <- cbind(rep(1:n.lati, n.long), rep(1:n.long, each=n.lati))
mat.dist <- as.matrix(dist(mat.lattice, upper=TRUE, diag=TRUE))
mat.cov <- sigmasq * rho^mat.dist

set.seed(1228)
# Generate regression (design) matrix with intercept
mat.X <- cbind(rep(1, n.site),scale(matrix(rnorm(n.site*(length(vec.beta)-1)),nrow=n.site)))
vec.Z <- t(chol(mat.cov)) %*% rnorm(n.site) + mat.X %*% vec.beta
vec.epsilon <- diag(sqrt(1-sigmasq), n.site) %*% rnorm(n.site)
vec.ylat <- as.numeric(vec.Z + vec.epsilon)

# Convert to the vector of observation
vec.yobs <- func.obs.ord(vec.ylat, vec.alpha=c(-Inf,0,vec.cutoff,Inf))
```

---

func.obs.prop *Latent Response Transformation for Proportional Data*

Description

func.obs.prop transforms a vector of latent responses into the corresponding observed ones under the spatial Tobit model.

Usage

```
func.obs.prop(vec.ylat, alpha)
```

Arguments

`vec.ylat`	a vector of latent responses for all N sites.
`alpha`	a cutoff point controlling the probability of latent reponse being one.

Value

`func.obs.prop` returns a vector of observed responses.

References

Feng, Xiaoping, Zhu, Jun, Lin, Pei-Sheng, and Steen-Adams, Michelle M. (2014) Composite likelihood Estimation for Models of Spatial Ordinal Data and Spatial Proportional Data with Zero/One values. *Environmetrics* 25(8): 571–583.

Examples

```
# A simple example for observation generation
a <- sample(c(0,1), 50, replace=TRUE)
b <- sample(runif(1000,0,10), 100, replace=TRUE)
alpha <- 4
vec.yobs <- func.obs.prop(vec.ylat=c(a, b), alpha=alpha)

# A complex example
# True parameter
alpha <- 4; vec.beta <- c(1, 2, 1, 0, -1); sigmasq <- 0.8; rho <- 0.6; radius <- 5
vec.par <- c(alpha, vec.beta, sigmasq, rho)

# Coordinate matrix
n.lati <- 30; n.long <- 30
n.site <- n.lati * n.long
mat.lattice <- cbind(rep(1:n.lati, n.long), rep(1:n.long, each=n.lati))
mat.dist <- as.matrix(dist(mat.lattice, upper=TRUE, diag=TRUE))
mat.cov <- sigmasq * rho^mat.dist

set.seed(1228)

# Generate regression (design) matrix with intercept
mat.X <- cbind(rep(1, n.site),scale(matrix(rnorm(n.site*(length(vec.beta)-1)),nrow=n.site)))
vec.Z <- t(chol(mat.cov)) %*% rnorm(n.site) + mat.X %*% vec.beta
vec.epsilon <- diag(sqrt(1-sigmasq), n.site) %*% rnorm(n.site)
vec.ylat <- as.numeric(vec.Z + vec.epsilon)

# Convert to the vector of observation
vec.yobs <- func.obs.prop(vec.ylat, alpha=alpha)
```

Index