

Package ‘SDALGCP’

August 14, 2018

Title Spatially Discrete Approximation to Log-Gaussian Cox Processes
for Aggregated Disease Count Data

Version 0.1.0

Description Provides a computationally efficient discrete approximation to log-Gaussian Cox process model for spatially aggregated disease count data. It uses Monte Carlo Maximum Likelihood for model parameter estimation as proposed by Christensen (2004) <doi: 10.1198/106186004X2525> and delivers prediction of spatially discrete and continuous relative risk.

Depends R (>= 3.4.0)

License GPL-2 | GPL-3

Encoding UTF-8

LazyData true

Imports pdist(>= 1.2), Matrix(>= 1.2.14), PrevMap(>= 1.4.1), raster(>= 2.6.7), sp(>= 1.2.7), spatstat(>= 1.55.1), splancs(>= 2.1.40), methods(>= 3.4.4), maptools(>= 0.9.2), plyr(>= 1.8.4), pbapply(>= 1.3.4), progress(>= 1.1.2)

RoxygenNote 6.0.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Olatunji Johnson [aut, cre],
Emanuele Giorgi [aut],
Peter Diggle [aut]

Maintainer Olatunji Johnson <olatunjijohnson21111@gmail.com>

Repository CRAN

Date/Publication 2018-08-14 09:30:05 UTC

R topics documented:

confint.SDALGCP	2
controlmcmcSDA	3

PBCshp	4
phiCI	5
plot.Pred.SDALGCP	6
pop_den	7
SDALGCP	8
SDALGCPMCMCML	9
SDALGCPPred	11
summary.SDALGCP	13

Index 15

confint.SDALGCP	<i>Confidence Intervals for SDALGCP Model Parameters</i>
-----------------	--

Description

Computes confidence intervals for one or more parameters in a fitted SDALGCP model from the object of class "SDALGCP", based on asymptotic normality.

Usage

```
## S3 method for class 'SDALGCP'
confint(object, parm, level = 0.95, dp = 3, ...)
```

Arguments

object	an object of class "SDALGCP" obtained as result of a call to SDALGCPMCMCML .
parm	a specification of which parameters are to be given confidence intervals, either a vector of numbers or a vector of names. If missing, all parameters are considered.
level	the confidence level required.
dp	the number of decimal places for the result
...	additional argument(s) for methods.

Value

A matrix (or vector) with columns giving lower and upper confidence limits for each parameter. These will be labelled as $(1-level)/2$ and $1 - (1-level)/2$ in

Author(s)

Olatunji O. Johnson <o.johnson@lancaster.ac.uk>
 Emanuele Giorgi <e.giorgi@lancaster.ac.uk>
 Peter J. Diggle <p.diggle@lancaster.ac.uk>

See Also

[confint.lm](#), [confint.default](#), [SDALGCPMCMCML](#)

controlmcmcSDA	<i>control.mcmcSDA</i>
----------------	------------------------

Description

This function helps to define the number of iteration, burn-in, thinning, and the tuning parameters of the adaptive MALA

Usage

```
controlmcmcSDA(n.sim, burnin, thin, h, c1.h, c2.h)
```

Arguments

n.sim	the number of iteration
burnin	The number of burn-in
thin	the number of thinning
h	tuning parameter of the proposal distribution used in the Langevin-Hastings MCMC algorithm (see Laplace.sampling); default is h=NULL and then set internally as $1.65/n(1/6)$, where n is the dimension of the random effect.
c1.h	value of c1 used in the adaptive scheme for h; default is c1.h=0.01. See also 'Details' in PrevMap package
c2.h	value of c2 used in the adaptive scheme for h; default is c2.h=1e-04. See also 'Details' in PrevMap package

Details

To be used in SDA

Value

A list with processed arguments to be passed to the main function.

Author(s)

Olatunji O. Johnson <o.johnson@lancaster.ac.uk>

Emanuele Giorgi <e.giorgi@lancaster.ac.uk>

Peter J. Diggle <p.diggle@lancaster.ac.uk>

See Also

[control.mcmc.MCML](#)

Examples

```
n <- 545
h <- 1.65/(n^(1/6))
control.mcmc <- controlmcmcSDA(n.sim = 10000, burnin = 2000,
thin= 8, h=h, c1.h = 0.01, c2.h = 1e-04)
str(control.mcmc)
```

PBCshp

PBC count data and index of multiple deprivation data.

Description

A dataset containing PBC count and Index of multiple deprivation

Usage

```
data(PBCshp)
```

Format

A SpatialPolygonsDataFrame of object containing the PBC cases count for each LSOA in Newcastle upon Tyne, UK, as well as the index of multiple deprivation.

X PBC count

pop population count

LSOA04CD LSOA ID

pop population count

males number of males

females number of females

propmale proportion of males

IMD index of multiple deprivation score

Income proportion of the population experiencing income deprivation

Employment proportion of the population experiencing employment deprivation

Health deprivation due to Health

Education deprivation due to education

Barriers barriers to housing and services

Crime deprivation due to crime

Environment living environment deprivation ...

References

Taylor, B., Davies, T., Rowlingson, B., & Diggle, P. (2015). Bayesian inference and data augmentation schemes for spatial, spatiotemporal and multivariate log-Gaussian Cox processes in R. *Journal of Statistical Software*, 63, 1-48.

phiCI	<i>Plot of the deviance to derive the confidence interval of of the scale parameter, phi</i>
-------	--

Description

This function computes the confidence interval of phi

Usage

```
phiCI(obj, coverage = 0.95, plot = TRUE)
```

Arguments

obj	object of class "SDALGCP" from the the call to function SDALGCPMCMCML
coverage	the coverage probability, default is 0.95
plot	logical, to plot the deviance curve. default is TRUE

Details

This function computes the confidence interval of phi

Value

return the deviance plot and the corresponding confidence interval of the scale parameter phi

Author(s)

Olatunji O. Johnson <o.johnson@lancaster.ac.uk>

Emanuele Giorgi <e.giorgi@lancaster.ac.uk>

Peter J. Diggle <p.diggle@lancaster.ac.uk>

See Also

[SDALGCPMCMCML](#)

plot.Pred.SDALGCP *plot.Pred.SDALGCP function*

Description

Simple plotting function for both discrete and continuous prediction from the object of class "Pred.SDALGCP".

Usage

```
## S3 method for class 'Pred.SDALGCP'
plot(x, type = "relrisk", continuous = NULL,
     thresholds = NULL, bound = NULL, ...)
```

Arguments

x	an object of class "Pred.SDALGCP" obtained as result of a call to SDALGCPPred .
type	Character string: what type of plot to produce. For discrete inference choices are "incidence" (=exp(mu+S)); "SEincidence" (standard error of incidence); "CovAdjRelRisk" (=exp(S)); or "SECovAdjRelRisk" (standard error of covariate adjusted relative risk); while for continuous inference, choices are "relrisk" (=exp(S)); "SErelrisk" (standard error of the relative risk).
continuous	logical; TRUE for spatially continuous relative risk and FALSE for region specific relative risk. default is TRUE
thresholds	optional; (only used if you want to plot the exceedance probability) either a vector of numbers or a vector of single value.
bound	optional; it gives the boundary of the region, only useful when the predictive location is supplied in SDALGCPPred
...	further arguments passed to plot .

Details

This function plots the inference from [SDALGCPPred](#) function. It plots for region-specific inference; incidence and covariate adjusted relative risk while for spatially continuous inference it plots the relative risk. It can as well plot the exceedance probability for spatially discrete and continuous inference.

Value

The function does not return any value.

Author(s)

Olatunji O. Johnson <o.johnson@lancaster.ac.uk>

Emanuele Giorgi <e.giorgi@lancaster.ac.uk>

Peter J. Diggle <p.diggle@lancaster.ac.uk>

See Also

[SDALGCPPred](#), [plot_continuous](#), [plot_discrete](#), [plot_SDALGCPexceedance](#), [SDALGCPexceedance](#)

Examples

```
### Prepare the input of the model
data(PBCshp)
data <- as.data.frame(PBCshp@data) #get the data
### Write the formula of the model
FORM <- X ~ propmale + Income + Employment + Education + Barriers + Crime +
Environment + offset(log(pop))
### set the discretised phi
phi <- seq(500, 1700, length.out = 20)
#### get the initial parameter
model <- glm(formula=FORM, family="poisson", data=data)
beta.start <- coef(model)
sigma2.start <- mean(model$residuals^2)
phi.start <- median(phi)
par0 <- c(beta.start, sigma2.start, phi.start)
# setup the control arguments for the MCMC
n <- 545
h <- 1.65/(n^(1/6))
control.mcmc <- controlmcmcSDA(n.sim = 10000, burnin = 2000,
                              thin= 8, h=h, c1.h = 0.01, c2.h = 1e-04)
###Run the model

my_est <- SDALGCPMCMC(formula=FORM, data=data, my_shp=PBCshp, delta=100, phi=phi, method=1,
                      weighted=FALSE, plot=TRUE, par0=NULL, control.mcmc=control.mcmc)
Con_pred <- SDALGCPPred(para_est=my_est, cellsize=300, continuous=TRUE)
#to plot the spatially continuous relative risk
plot(Con_pred, type="relrisk")
#to plot the incidence
plot(Con_pred, type="incidence", continuous=FALSE)
#to plot the exceedance probability of the relative risk
plot(Con_pred, type="relrisk", thresholds= 2)
#to plot the exceedance probability of the incidence
plot(Con_pred, type="incidence", continuous=FALSE, thresholds= 0.001)
```

pop_den

Population density raster shapefile of Newcastle Upon Tyne in the UK.

Description

A dataset containing the population density of Newcastle upon Tyne of 300 X 300 metres

Usage

```
data(pop_den)
```

Format

This is a raster file of population density, 300 X 300 meters.

References

Murdock, A.P., Harfoot, A.J.P., Martin, D., Cockings, S. and Hill, C. (2015) OpenPopGrid: an open gridded population dataset for England and Wales. GeoData, University of Southampton.

SDALGCP

SDALGCP: A package to make continuous inference from spatially aggregated disease count data.

Description

The SDALGCP package provides two categories of important functions: SDALGCPMML and SDALGCPPred.

SDALGCP functions

The [SDALGCPMML](#) function uses Monte Carlo Maximum Likelihood to estimate the parameter of a poisson log-linear model with spatially continuous random effect.

The [SDALGCPPred](#) function delivers spatially discrete prediction of the incidence and the covariate adjusted relative risk and spatially continuous prediction of the covariate adjusted relative risk.

Functions such as [summary](#), [confint](#) and [print](#) also can be applied to the output.

Author(s)

Olatunji O. Johnson, Emanuele Giorgi, Peter Diggle. All from CHICAS, Lancaster Medical School, Faculty of Health and Medicine, Lancaster University

References

Christensen, O. F. (2004). Monte carlo maximum likelihood in model-based geostatistics. *Journal of Computational and Graphical Statistics* 13, 702-718.

Giorgi, E., & Diggle, P. J. (2017). PrevMap: an R package for prevalence mapping. *Journal of Statistical Software*, 78(8), 1-29. doi:10.18637/jss.v078.i08

Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2014). *Hierarchical modeling and analysis for spatial data*. CRC press.

SDALGCPMCMCML

*Parameter estimation for SDA-LGCP Using Monte Carlo Maximum likelihood***Description**

This function provides the maximum likelihood estimation of the parameter given a set of values of scale parameter of the Gaussian process, phi.

Usage

```
SDALGCPMCMCML(formula, data, my_shp, delta, phi = NULL, method = 1,
  pop_shp = NULL, weighted = FALSE, par0 = NULL, control.mcmc = NULL,
  plot = FALSE, plot_profile = TRUE, rho = NULL, giveup = NULL,
  messages = FALSE)
```

Arguments

formula	an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted.
data	data frame containing the variables in the model.
my_shp	A <code>SpatialPolygons</code> or <code>SpatialPolygonsDataFrame</code> object containing the polygons (i.e each regions).
delta	distance between points
phi	the discretised values of the scale parameter phi. if not supplied, it uses the default, which is 20 phis' which ranges from size of the smallest region to the one-tenth of the size of the entire domain.
method	To specify which method to use to sample the points, the options are 1 for Simple Sequential Inhibition (SSI) process, 2 for Uniform sampling and 3 for regular grid. 1 is the default
pop_shp	Optional, The raster of population density map for population weighted approach
weighted	To specify if you want to use the population density, default to FALSE, i.e population density is not used.
par0	the initial parameter of the fixed effects beta, the variance sigmasq and the scale parameter phi, specified as <code>c(beta, sigma2, phi)</code> . Default; beta, the estimates from the glm; sigma2, variance of the residual; phi, the median of the supplied phi.
control.mcmc	list from <code>PrevMap</code> package to define the burnin, thinning, the number of iteration and the turning parameters see controlmcmcSDA .
plot	To display the plot of the points inside the polygon, default to TRUE
plot_profile	logical; if TRUE the profile-likelihood is plotted. default is FALSE
rho	Optional, the packing density, default set to 0.55

giveup	Optional, number of rejected proposals after which the algorithm should terminate, default set to 1000
messages	logical; if messages=TRUE, it prints the results objective function and the parameters at every phi iteration. Default is FALSE.

Details

This function performs parameter estimation for a SDALGCP Model **Monte Carlo Maximum likelihood**. The Monte Carlo maximum likelihood method uses conditional simulation from the distribution of the random effect $T(x) = d(x)' \beta + S(x)$ given the data y , in order to approximate the high-dimensional intractable integral given by the likelihood function. The resulting approximation of the likelihood is then maximized by a numerical optimization algorithm which uses analytic expression for computation of the gradient vector and Hessian matrix. The functions used for numerical optimization are `nlminb`. The first stage of estimation is generating locations inside the polygon, followed by precomputing the correlation matrices, then optimising the likelihood.

Value

An object of class "PrevMap". The function `summary.PrevMap` is used to print a summary of the fitted model. The object is a list with the following components:

D: matrix of covariates.

y: the count, response observations.

m: offset

beta_opt: estimates of the fixed effects of the model.

sigma2_opt: estimates of the variance of the Gaussian process.

phi_opt: estimates of the scale parameter phi of the Gaussian process.

cov: covariance matrix of the MCMC estimates.

Sigma_mat_opt: covariance matrix of the Gaussian process that corresponds to the optimal value

llike_val_opt: maximum value of the log-likelihood.

mu: mean of the linear predictor

all_para: the entire estimates for the different values of phi.

all_cov: the entire covariance matrix of the estimates for the different values of phi.

par0: the initial parameter of the fixed effects beta and the variance sigmasq used in the estimation
 control.mcmc: the burnin, thinning, the number of iteration and the turning parameters used see [controlmcmcSDA](#).

call: the matched call.

Author(s)

Olatunji O. Johnson <o.johnson@lancaster.ac.uk>

Emanuele Giorgi <e.giorgi@lancaster.ac.uk>

Peter J. Diggle <p.diggle@lancaster.ac.uk>

References

- Giorgi, E., & Diggle, P. J. (2017). PrevMap: an R package for prevalence mapping. *Journal of Statistical Software*, 78(8), 1-29. doi:10.18637/jss.v078.i08
- Christensen, O. F. (2004). Monte Carlo maximum likelihood in model-based geostatistics. *Journal of Computational and Graphical Statistics* 13, 702-718.

See Also

[Aggregated_poisson_log_MCML](#), [Laplace.sampling](#), [summary.SDALGCP](#)

Examples

```
### Prepare the input of the model
data(PBCshp)
data <- as.data.frame(PBCshp@data) #get the data
### Write the formula of the model
FORM <- X ~ propmale + Income + Employment + Education + Barriers + Crime +
Environment + offset(log(pop))
### set the discretised phi
phi <- seq(500, 1700, length.out = 20)
#### get the initial parameter
model <- glm(formula=FORM, family="poisson", data=data)
beta.start <- coef(model)
sigma2.start <- mean(model$residuals^2)
phi.start <- median(phi)
par0 <- c(beta.start, sigma2.start, phi.start)
# setup the control arguments for the MCMC
n <- 545
h <- 1.65/(n^(1/6))
control.mcmc <- controlmcmcSDA(n.sim = 10000, burnin = 2000,
  thin= 8, h=h, c1.h = 0.01, c2.h = 1e-04)
###Run the model

my_est <- SDALGCPMCMC(formula=FORM, data=data, my_shp=PBCshp, delta=100, phi=phi, method=1,
  weighted=FALSE, plot=TRUE, par0=par0, control.mcmc=control.mcmc)
```

SDALGCPPred

*Spatial continuous predictions of the relative risk using plug-in of
MCML estimates*

Description

This function performs spatial continuous prediction, fixing the model parameters at the Monte Carlo maximum likelihood estimates of a SDALGCP model.

Usage

```
SDALGCPPred(para_est, cellsize, continuous = TRUE, control.mcmc = NULL,
  pred.loc = NULL, divisor = 1, plot.correlogram = F, messages = TRUE,
  parallel = FALSE)
```

Arguments

<code>para_est</code>	an object of class "SDALGCP" obtained as a result of a call to SDALGCPMCMC .
<code>cellsize</code>	the size of the computational grid
<code>continuous</code>	logical; to choose which prediction to do perform, discrete or continuous. the default is continuous.
<code>control.mcmc</code>	output from controlmcmcSDA , if not provided, it uses the values used for the parameter estimation
<code>pred.loc</code>	optional, the dataframe of the predictive grid.
<code>divisor</code>	optional, the value to use to convert the dimension of the polygon, default is 1 which implies no conversion
<code>plot.correlogram</code>	logical; if <code>plot.correlogram=TRUE</code> the autocorrelation plot of the conditional simulations is displayed.
<code>messages</code>	logical; if <code>messages=TRUE</code> then status messages are printed on the screen (or output device) while the function is running. Default is <code>messages=TRUE</code> .
<code>parallel</code>	to parallelize some part of the function.

Details

The function perform prediction of the spatially discrete incidence and covariate adjusted relative risk, and spatially continuous relative risk. The discrete inference uses the Metropolis-Adjusted Langevin Hasting sampling from [Laplace.sampling](#). And the continuous inference is typically change of support inference.

Value

`pred.draw`: the samples of the prediction
`pred`: the prediction of the relative risk
`predSD`: the standard error of the prediction
`Pred.loc`: The coordinates of the predictive locations

Author(s)

Olatunji O. Johnson <o.johnson@lancaster.ac.uk>
 Emanuele Giorgi <e.giorgi@lancaster.ac.uk>
 Peter J. Diggle <p.diggle@lancaster.ac.uk>

References

Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2014). Hierarchical modeling and analysis for spatial data. CRC press.

See Also

[plot.Pred.SDALGCP](#), [SDAContinuousPred](#), [SDADiscretePred](#), [plot_continuous](#), [plot_discrete](#)

Examples

```
### Prepare the input of the model
data(PBCshp)
data <- as.data.frame(PBCshp@data) #get the data
### Write the formula of the model
FORM <- X ~ propmale + Income + Employment + Education + Barriers + Crime +
Environment + offset(log(pop))
### set the discretised phi
phi <- seq(500, 1700, length.out = 20)
#### get the initial parameter
model <- glm(formula=FORM, family="poisson", data=data)
beta.start <- coef(model)
sigma2.start <- mean(model$residuals^2)
phi.start <- median(phi)
par0 <- c(beta.start, sigma2.start, phi.start)
# setup the control arguments for the MCMC
n <- 545
h <- 1.65/(n^(1/6))
control.mcmc <- controlmcmcSDA(n.sim = 10000, burnin = 2000,
                              thin= 8, h=h, c1.h = 0.01, c2.h = 1e-04)
###Run the model

my_est <- SDALGCPMCMC(formula=FORM, data=data, my_shp=PBCshp, delta=100, phi=phi, method=1,
                     weighted=FALSE, plot=TRUE, par0=par0, control.mcmc=control.mcmc)
Con_pred <- SDALGCPPred(para_est=my_est, cellsize=300, continuous=TRUE)
```

summary.SDALGCP

Summarizing the parameter estimates of SDALGCP model

Description

summary method for the class "SDALGCP" that computes the standard errors and p-values of SDALGCP.

Usage

```
## S3 method for class 'SDALGCP'
summary(object, ...)
```

Arguments

object an object of class "SDALGCP" obtained as result of a call to [SDALGCPMCMC](#) .
... further arguments passed to or from other methods.

Value

A list with the following components
parameter_estimate_result: the parameter of the SDALGCP model
phi: the scale parameter of the Gaussian process
ll: value of likelihood function at the maximum likelihood estimates.
call: matched call.

Author(s)

Olatunji O. Johnson <o.johnson@lancaster.ac.uk>
Emanuele Giorgi <e.giorgi@lancaster.ac.uk>
Peter J. Diggle <p.diggle@lancaster.ac.uk>

Index

*Topic **datasets**

PBCshp, [4](#)
pop_den, [7](#)

Aggregated_poisson_log_MCML, [11](#)

confint, [8](#)
confint.default, [2](#)
confint.lm, [2](#)
confint.SDALGCP, [2](#)
control.mcmc.MCML, [3](#)
controlmcmcSDA, [3](#), [9](#), [10](#), [12](#)

formula, [9](#)

Laplace.sampling, [11](#), [12](#)

nlminb, [10](#)

PBCshp, [4](#)
phiCI, [5](#)
plot, [6](#)
plot.Pred.SDALGCP, [6](#), [13](#)
plot_continuous, [7](#), [13](#)
plot_discrete, [7](#), [13](#)
plot_SDALGCPexceedance, [7](#)
pop_den, [7](#)
print, [8](#)

SDAContinuousPred, [13](#)
SDADiscretePred, [13](#)
SDALGCP, [8](#)
SDALGCP-package (SDALGCP), [8](#)
SDALGCPexceedance, [7](#)
SDALGCPMCML, [2](#), [5](#), [8](#), [9](#), [12](#), [14](#)
SDALGCPPred, [6–8](#), [11](#)
summary, [8](#)
summary.PrevMap, [10](#)
summary.SDALGCP, [11](#), [13](#)