

Package ‘scp’

May 8, 2020

Type Package

Title Spatial Conformal Prediction

Version 0.1.0

Author Huiying Mao <hmao@samsi.info>, Brian Reich <bjreich@ncsu.edu>, Ryan Martin <rgmarti3@ncsu.edu>

Maintainer Huiying Mao <hmao@samsi.info>

Description Provide robust, valid, and model-free spatial prediction intervals using Spatial Conformal Prediction (SCP) algorithms

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

URL <https://github.com/mhuiying/scp>

BugReports <https://github.com/mhuiying/scp/issues>

Suggests knitr,
rmarkdown,
covr,
testthat

VignetteBuilder knitr

Imports geoR

R topics documented:

krige_pred	2
plausibility	3
plausibility_contour	4
scp	6

Index	9
--------------	----------

krige_pred	<i>Kriging prediction function</i>
------------	------------------------------------

Description

This function provides an example for `pred_fun` in [scp](#), [plausibility](#), and [plausibility.contour](#), which provides a point prediction for location `s0` (and corresponding standard error), given observations `s` and `Y`.

Usage

```
krige_pred(s0, s, Y, return_sd = FALSE)
```

Arguments

<code>s0</code>	prediction location, a numeric vector with <code>length = 2</code> .
<code>s</code>	an $n \times 2$ matrix or <code>data.frame</code> with two coordinates of n locations.
<code>Y</code>	a vector with n values corresponding to $Y(s)$.
<code>return_sd</code>	logical. if TRUE, <code>Krige_pred</code> function returns the standard error of $Y(s0)$ along with the point prediction; if FALSE, <code>Krige_pred</code> function only returns the point prediction. Defaults to FALSE.

Value

If `return_sd` is FALSE (default), the output is a value of point prediction for $Y(s0)$; If `return_sd` is TRUE, the output is a list with the following elements:

<code>yhat</code>	point prediction for $Y(s0)$
<code>sd</code>	standard error for $Y(s0)$

Examples

```
#?sample_data
s0 = c(0.5,0.5)
s = sample_data$s
Y = sample_data$Y

krige_pred(s0,s,Y)
krige_pred(s0,s,Y,return_sd=TRUE)
```

plausibility	<i>calculate plausibility of $Y(s_0)$ being Y_0</i>
--------------	---

Description

This function provides the plausibility of $Y(s_0)$ being Y_0 , given observations s and Y , using spatial conformal prediction algorithms.

Usage

```
plausibility(
  Y0,
  s0,
  s,
  Y,
  global = TRUE,
  eta = Inf,
  m = NULL,
  pred_fun = krige_pred,
  thetaHat = NULL,
  dfun = c("residual2", "std_residual2")
)
```

Arguments

Y0	a scalar or a vector
s0	prediction location, a numeric vector with <code>length = 2</code> or a <code>matrix</code> with 1 row and 2 cols, or a <code>data.frame</code> with 1 row and 2 coordinates.
s	an $n \times 2$ <code>matrix</code> or a <code>data.frame</code> with two coordinates of n locations.
Y	a vector with n values corresponding to $Y(s)$.
global	logical; if <code>TRUE</code> , <code>scp</code> function returns the result of global spatial conformal prediction (GSCP); if <code>FALSE</code> , <code>scp</code> function returns the result of local spatial conformal prediction (LSCP) and users need to specify <code>eta < Inf</code> or <code>m ≤ n</code> . Defaults to <code>TRUE</code> .
eta	kernel bandwidth for weight schema, a positive scalar with smaller value meaning more localized procedure. Defaults to <code>Inf</code> , which puts equal weight on surrounding <code>m</code> points.
m	an positive integer representing the number of nearest locations to use for prediction. Default to <code>NULL</code> . If <code>global = TRUE</code> , <code>m = n</code> ; if <code>global = FALSE</code> and <code>m</code> is not specified, <code>m</code> would be determined by <code>eta</code> .
pred_fun	spatial prediction function with inputs being <code>s0</code> , <code>s</code> , <code>Y</code> and ouputs being predicted $Y(s_0)$ (and its standard error). Defaults to krige_pred .
thetaHat	a vector of Matern parameters, representing nugget, partial sill, range, and smoothness as in Mao. et al. (2020). Defaults to <code>NULL</code> . It will be ignored if <code>pred_fun</code> is not <code>krige_pred</code> .

dfun non-conformity measure with four options. In which, "residual2" (default) represents squared residual and "std_residual2" represents standardized squared residual.

Value

The output is a scalar or a vector with plausibility values for Y_0 . The numbers are between 0 and 1.

Author(s)

Huiying Mao, <hmao@samsi.info>, Brian Reich <bjreich@ncsu.edu>

References

to be entered

See Also

[plausibility_contour](#)

Examples

```
## To predict  $Y(s_0)$ , where  $s_0 = c(0.5, 0.5)$ , using sample data
## What's the plausibility if  $Y(s_0) = 0$ ?  $Y(s_0) = 1.5$ ?

#?sample_data
s0 = c(0.5, 0.5)
s = sample_data$s
Y = sample_data$Y

# plausibility for  $Y(s_0) = 0$ 
plausibility(Y0=0, s0=s0, s=s, Y=Y)

# plausibility for  $Y(s_0) = 1.5$ 
plausibility(Y0=1.5, s0=s0, s=s, Y=Y)

# plausibility for a sequence of  $Y_0$ 's
plausibility(Y0=seq(0, 1, 0.1), s0=s0, s=s, Y=Y)
```

plausibility_contour *generate plausibility contour*

Description

This function provides the plausibility contour for $Y(s_0)$, given observations s and Y , using spatial conformal prediction algorithms.

Usage

```
plausibility_contour(
  s0,
  s,
  Y,
  global = TRUE,
  eta = Inf,
  m = NULL,
  pred_fun = krige_pred,
  thetaHat = NULL,
  dfun = c("residual2", "std_residual2"),
  precision = NULL
)
```

Arguments

<code>s0</code>	prediction location, a numeric vector with <code>length = 2</code> or a <code>matrix</code> with 1 row and 2 cols, or a <code>data.frame</code> with 1 row and 2 coordinates.
<code>s</code>	an $n \times 2$ <code>matrix</code> or a <code>data.frame</code> with two coordinates of n locations.
<code>Y</code>	a vector with n values corresponding to $Y(s)$.
<code>global</code>	logical; if <code>TRUE</code> , <code>scp</code> function returns the result of global spatial conformal prediction (GSCP); if <code>FALSE</code> , <code>scp</code> function returns the result of local spatial conformal prediction (LSCP) and users need to specify <code>eta < Inf</code> or <code>m ≤ n</code> . Defaults to <code>TRUE</code> .
<code>eta</code>	kernel bandwidth for weight schema, a positive scalar with smaller value meaning more localized procedure. Defaults to <code>Inf</code> , which puts equal weight on surrounding <code>m</code> points.
<code>m</code>	an positive integer representing the number of nearest locations to use for prediction. Default to <code>NULL</code> . If <code>global = TRUE</code> , <code>m = n</code> ; if <code>global = FALSE</code> and <code>m</code> is not specified, <code>m</code> would be determined by <code>eta</code> .
<code>pred_fun</code>	spatial prediction function with inputs being <code>s0, s, Y</code> and outputs being predicted $Y(s0)$ (and its standard error). Defaults to <code>krige_pred</code> .
<code>thetaHat</code>	a vector of Matern parameters, representing nugget, partial sill, range, and smoothness as in Mao. et al. (2020). Defaults to <code>NULL</code> . It will be ignored if <code>pred_fun</code> is not <code>krige_pred</code> .
<code>dfun</code>	non-conformity measure with four options. In which, <code>"residual2"</code> (default) represents squared residual and <code>"std_residual2"</code> represents standardized squared residual.
<code>precision</code>	a positive scalar represents how dense $Y(s)$ candidates (<code>Y_cand</code>) are. Defaults to <code>NULL</code> .

Value

The output is a “`plausibility_contour`” object.

Author(s)

Huiying Mao, <hmao@samsi.info>, Brian Reich <bjreich@ncsu.edu>

References

to be entered

See Also

[plausibility](#)

Examples

```
## generate plausibility contour for Y(s0), where s0 = c(0.5,0.5), using sample data

#?sample_data
s0 = c(0.5,0.5)
s = sample_data$s
Y = sample_data$Y

pc = plausibility_contour(s0=s0,s=s,Y=Y)
plot(pc)

idx = which(s[, 1] == s0[1] & s[, 2] == s0[2])
abline(v = Y[idx], col = "red", lty = 2)
legend("topright", col=1:2, lty=1:2, c("plausibility", "true value"))
```

 scp

Spatial conformal prediction at input location(s)

Description

This function provides the spatial conformal prediction interval for location(s) `s0`, given observations `s` and `Y`.

Usage

```
scp(
  s0,
  s,
  Y,
  global = TRUE,
  eta = Inf,
  m = NULL,
  pred_fun = krige_pred,
  thetaHat = NULL,
  dfun = c("residual2", "std_residual2"),
```

```

    precision = NULL,
    alpha = 0.05
  )

```

Arguments

<code>s0</code>	prediction location(s), a numeric vector with <code>length = 2</code> , or a <code>matrix</code> with <code>ncol = 2</code> , or a <code>data.frame</code> with two coordinates.
<code>s</code>	an $n \times 2$ <code>matrix</code> or a <code>data.frame</code> with two coordinates of n locations.
<code>Y</code>	a vector with n values corresponding to $Y(s)$.
<code>global</code>	logical; if <code>TRUE</code> , <code>scp</code> function returns the result of global spatial conformal prediction (GSCP); if <code>FALSE</code> , <code>scp</code> function returns the result of local spatial conformal prediction (LSCP) and users need to specify <code>eta</code> . Defaults to <code>TRUE</code> .
<code>eta</code>	kernel bandwidth for weight schema, a positive scalar with smaller value meaning more localized procedure. Defaults to <code>Inf</code> , which puts equal weight on surrounding m points.
<code>m</code>	an positive integer representing the number of nearest locations to use for prediction. Default depends on <code>eta</code> .
<code>pred_fun</code>	spatial prediction function with inputs being <code>s0</code> , <code>s</code> , <code>Y</code> and outputs being predicted $Y(s0)$ (and its standard error). Defaults to <code>krige_pred</code> representing Kriging prediction.
<code>thetaHat</code>	a vector of Matern parameters, representing nugget, partial sill, range, and smoothness as in Mao. et al. (2020). Defaults to <code>NULL</code> . It will be ignored if <code>pred_fun</code> is not <code>krige_pred</code> .
<code>dfun</code>	non-conformity measure with four options. In which, <code>"residual2"</code> (default) represents squared residual and <code>"std_residual2"</code> represents standardized squared residual.
<code>precision</code>	a positive scalar represents how dense the candidates for $Y(s)$ are. Defaults to <code>NULL</code> .
<code>alpha</code>	significance level. Defaults to 0.05.

Value

The output is a `data.frame` of lower and upper bounds of the conformal prediction interval(s) for `s0`.

Author(s)

Huiying Mao, <hmao@samsi.info>, Brian Reich <bjreich@ncsu.edu>

References

to be entered

See Also

[plausibility](#), [plausibility_contour](#)

Examples

```
## generate prediction interval for prediction locations(s) s0(s) using sample data

#?sample_data
s = sample_data$s
Y = sample_data$Y

# locations to predict
s0 = c(0.5,0.5)
s0s = rbind(c(0.4, 0.4), c(0.5,0.5), c(0.6, 0.6))

# default prediction interval
scp(s0=s0,s=s,Y=Y)
scp(s0=s0s,s=s,Y=Y)

# user define eta=0.1, where LSCP is considered
scp(s0=s0,s=s,Y=Y,eta=0.1)

# user define non-conformity measure
scp(s0=s0,s=s,Y=Y,dfun="std_residual2")

# user define prediction function
fun = function(s0,s,Y) return(mean(Y))
scp(s0=s0,s=s,Y=Y,pred_fun=fun)
```


Index

krige_pred, [2](#), [3](#), [5](#), [7](#)

plausibility, [2](#), [3](#), [6](#), [7](#)

plausibility_contour, [2](#), [4](#), [4](#), [7](#)

scp, [2](#), [6](#)