Package 'OptimaRegion'

May 20, 2016

Type Package

Title Confidence Regions for Optima

Version 0.2

Author Enrique del Castillo, John Hunt, and James Rapkin

Maintainer Enrique del Castillo <exd13@psu.edu>

Depends R (>= 3.2.3)

Imports fields,rsm,nloptr,DepthProc,boot,spam

Description Computes confidence regions on the location of response surface optima.

License GPL (>= 2)

LazyData TRUE

RoxygenNote 5.0.1

NeedsCompilation no

R topics documented:

`OptimaRegion-package` *OptimaRegion*

Description

OptimaRegion is a package for the computation of confidence regions on the location of the optima of response surface models. Both parametric (quadratic polynomial) and nonparametric (thin plate spline) models are supported. The confidence regions obtained do not rely on any distributional assumption, such as Normality of the response, and are obtained by bootstrapping. The resulting regions are both valid and unbiased, and have a size that rapidly decreases as the sample size

increases. Regions are obtained both numerically (as a set of points) and graphically, as the convex hull of the points. Functionality for the computation of a bootstrap confidence interval on the distance between the optima of two different response surfaces is included.

Author(s)

Enrique del Castillo <exd13@psu.edu>, John Hunt <J.Hunt@exeter.ac.uk> and James Rapkin <jr297@exeter.ac.uk>

References

Del Castillo, E., Hunt, J., Rapkin, J., and Zarmehri, S. , "Confidence regions for the location of response surface optima: the R package OptimaRegion".

`CRcompare` *Computes bootstrapped confidence intervals for the distance between two response surface optima*

Description

Computes bootstrapped confidence intervals for the mean and median distance between the optima of two response surface models. Models can be thin plate splines or quadratic polynomials.

Usage

```
CRcompare(X1,y1,X2,y2,lambda=0.04,responseType='TPS',nosim1and2=200,
alpha=0.05,LB1,UB1,triangularRegion1=FALSE, vertex11=NULL, vertex21=NULL,
maximization1=TRUE,outputPDFFile1="CR_plot1.pdf",outputOptimafile1="Optima1.txt"
,LB2,UB2,triangularRegion2=FALSE, vertex12=NULL, vertex22=NULL, maximization2=
TRUE,outputPDFFile2="CR_plot2.pdf",outputOptimafile2="Optima2.txt",xlab1and2=
"Protein eaten (mg)",ylab1and2="Carbohydrates eaten (mg)")
```

Arguments

`X1`	nx2 matrix with the values of the 2 regressors (experimental factors) in the n observations associated with response 1
`y1`	nx1 vector of response no. 1 value observations
`X2`	nx2 matrix with the values of the 2 regressors (experimental factors) in the n observations associated with response 2
`y2`	nx1 vector of response no. 2 value observations
`lambda`	smoothing penalty if a TPS model is selected (default=0.04)
`responseType`	use 'TPS' if fitting thin plate spline responses, 'Quad' if fitting quadratic polynomials
`nosim1and2`	number of simulations (default=200)
`alpha`	confidence level (0<alpha<1; default=0.05)
`LB1,UB1`	2x1 vectors of lower and upper bounds for search region where optima may lie for response 1

`triangularRegion1`
: logical: if TRUE it will constrain the optimum points of response 1 to lie inside a triangle defined by the coordinates (0,0), and those in "vertex11", and "vertex21", see below). NOTE: use TRUE when the treatments form a triangular experimental region in shape. If FALSE, maxima will only be constrained to lie inside the rectangular region defined by LB1 and UB1. Default is FALSE.

`vertex11,vertex21`
: 2x1 vectors with coordinates defining two of the 3 vertices of a triangular region for searching the optima of response 1. Must be provided if triangularRegion is TRUE (NOTE: vertices numbered clockwise, vertex0=c(0,0) always)

`maximization1`
: logical: if TRUE (default) it maximizes response 1, if FALSE it minimizes it

`outputPDFFile1`
: name of the PDF file where the CR plot of response 1 is saved (default: "CRplot1.pdf")

`outputOptimafile1`
: name of text file for saving the coordinates of the optima of response 1 (default: "Optima1.txt")

`LB2,UB2`
: 2x1 vectors of lower and upper bounds for search region where optima may lie for response 2

`triangularRegion2`
: logical: if TRUE it will constrain the optimum points of response 2 to lie inside a triangle defined by the coordinates (0,0), and those in "vertex12", and "vertex22", see below (in addition to being constrained to lie inside the region defined by LB and UB). NOTE: use TRUE when the treatments form a triangular experimental region in shape. If FALSE, maxima will only be constrained to lie inside the rectangular region defined by LB2 and UB2. Default is FALSE.

`vertex12`
: see vertex22

`vertex22`
: 2x1 vectors with coordinates defining two of the 3 vertices of a triangular region for searching the optima of response 2. Must be provided if triangularRegion is TRUE (NOTE: vertices numbered clockwise, vertex0=c(0,0) always)

`maximization2`
: logical: if TRUE (default) it maximizes response 2 if FALSE it minimizes it

`outputPDFFile2`
: name of the PDF file where the CR plot of response 2 is saved (default: "CRplot2.pdf")

`outputOptimafile2`
: name of text file for saving the coordinates of the optima of response 2 (default: "Optima2.txt")

`xlab1and2`
: text label for x axis in confidence region plot for the optima of each response (default: "Protein eaten, mg")

`ylab1and2`
: text label for y axis in confidence region plot for the optima of each response (default: "Carbohydrates eaten, mg")

Details

Computes distribution-free bootstrapped confidence intervals on the mean and median distance between the optima of two different responses. The responses can be Thin Plate Spline models or Quadratic polynomial models. Program calls OptRegionTps.R or OptRegionQuad.R to compute confidence regions on the optima of each response, next computes all pairwise distances between points in each CR, and finally bootstraps the distances to compute bca bootstrapped confidence intervals for the mean and median distance.

Usage assuming all default options:

out<-CRcompare(X1=X1,y1=y1,X2=X2,y2=y2,LB1=LB1,UB1=UB1,LB2=LB2,UB2=UB2)

Value

dist	vector of distances between pairs of points taken from each set of optima
mean	mean of all pairwise distances
median	median of all pairwise distances
ciMean	(1-alpha)*100% confidence interval for the mean of the distances using bca bootstrapping
ciMedian	(1-alpha)*100% confidence interval for the mean of the distances using bca bootstrapping. Note: ciMean and ciMedian are vectors with 5 columns, containing the significance level, the next two containing the indices of the order statistics used in the calculations and the final two the calculated endpoints of the CI's.

Note

Upon completion, PDF files containing the CR plots of the optima of each response with name as set in ouputPDFFile1 and outPDFFile2 are created. Two text files named as set in outputOptimaFile1 and outputOptimaFile2 are also created. These contain all the individual optima of each response in each CR.

Author(s)

Enrique del Castillo <exd13@psu.edu>, John Hunt <J.Hunt@exeter.ac.uk> and James Rapkin <jr297@exeter.ac.uk>

References

Del Castillo, E., Hunt, J., Rapkin, J., and Zarmehri, S. , "Confidence regions for the location of response surface optima: the R package OptimaRegion".

Examples

```
## Example: two randomly generated data sets, quadratic polynomial responses.

X1<-cbind(runif(100,-2,2),runif(100,-2,2))

y1<-as.matrix(72-11.78*X1[,1]+0.74*X1[,2]-7.25*X1[,1]^2-7.55*X1[,2]^2-4.85*X1[,1]*X1[,2]+
                  rnorm(100,0,8))

X2<-cbind(runif(100,-2,2),runif(100,-2,2))

y2<-as.matrix(72-11.78*X2[,1]+0.74*X2[,2]-7.25*X2[,1]^2-7.55*X2[,2]^2-4.85*X2[,1]*X2[,2]+
                  rnorm(100,0,8))

out<-CRcompare(X1=X1,y1=y1,X2=X2,y2=y2,responseType='Quad',nosim1and2=200,
alpha=0.05,LB1=c(-2,-2),UB1=c(2,2),LB2=c(-2,-2),UB2=c(2,2))
```

Drug *Mixture-amount experiment*

Description

A pharmaceutical mixture-amount experiment in two components

Usage

```
data("Drug")
```

Format

A data frame with 360 observations on the following 3 variables.

'Component_1' Component 1 amount (mg)

'Component_2' Component 2 amount (mg)

Percent Percent of cells killed (response)

Examples

```
plot(Drug[,1:2])
```

OptRegionQuad *Computes Confidence Regions of Optima of Quadratic Polynomial Models*

Description

Computes and displays the confidence region on the location of the optima of a quadratic response surface in 2 factors using bootstrapping.

Usage

```
OptRegionQuad(X, y, nosim=200, alpha=0.05, LB, UB,
triangularRegion=FALSE, vertex1=NULL, vertex2=NULL,
maximization=TRUE,
xlab="Protein eaten, mg",
ylab="Carbohydrates eaten, mg",
outputPDFFile="CRplot.pdf")
```

Arguments

X	nx2 matrix with the values of the 2 regressors (experimental factors) in the n observations
y	nx1 vector of response value observations
nosim	number of simulations (default=200)
alpha	confidence level (0<alpha<1; default=0.05)
LB,UB	2x1 vectors of lower and upper bounds for search region where optima may lie
triangularRegion	logical: if TRUE it will constrain the optimum points to lie inside a triangle defined by the coordinates (0,0), and those in "vertex1", and "vertex2", see below (in addition to being constrained to lie inside the region defined by LB and UB). NOTE: use TRUE when the treatments form a triangular experimental region in shape. If FALSE, maxima will only be constrained to lie inside the rectangular region defined by LB and UB. Default is FALSE.
vertex1,vertex2	2x1 vectors with coordinates defining two of the 3 vertices of a triangular region. Must be provided if triangularRegion is TRUE (NOTE: vertices numbered clockwise, vertex0=c(0,0) always)
maximization	logical: if TRUE (default) it maximizes it FALSE it minimizes
xlab	text label for x axis in confidence region plot (default: "Protein eaten, mg")
ylab	text label for y axis in confidence region plot (default: "Carbohydrates eaten, mg")
outputPDFFile	name of the PDF file where the CR plot is saved (default: "CRplot.pdf")

Details

Computes and displays an approximated 100*(1-alpha) percent confidence region (CR) for the linear-constrained maximum of a quadratic polynomial regression model in 2 controllable factors. Grey region on output plot is the approximate CR. The CR is computed as the convex hull of the coordinates of the optima found from bootstrapping nosim quadratic polynomial regressions to the data (therefore, it is an approximate CR). The mean value of the optimum is shown as a red point, and a smoothed contour plot of the X,y data obtained via thin plate splines is shown as well.

Usage assuming all default options:

out<-OptRegionQuad(X=X,y=y,LB=LB,UB=UB)

Value

meanPoint	a 2x1 vector with the coordinates of the mean optimum point
xin	an mx2 matrix with the x,y coordinates of all simulated #points that belong to the confidence region (dim(m) is (1-alpha)*nosim)

Note

Upon completion, a PDF file containing the CR plot with name as set in ouputPDFFile is created. To help visualization of the CR and the experimental response data, the plot also shows contours of a smoothed approximation to the response values, note this is not the quadratic polynomial that was fit (it is actually a TPS model).

Author(s)

Enrique del Castillo <exd13@psu.edu>, John Hunt <J.Hunt@exeter.ac.uk> and James Rapkin <jr297@exeter.ac.uk>

References

Del Castillo, E., Hunt, J., Rapkin, J., and Zarmehri, S. , "Confidence regions for the location of response surface optima: the R package OptimaRegion".

Examples

```
## Example 1: randomly generated 2-variable response surface data

X<-cbind(runif(100,-2,2),runif(100,-2,2))

y<-as.matrix(72-11.78*X[,1]+0.74*X[,2]-7.25*X[,1]^2-7.55*X[,2]^2-4.85*X[,1]*X[,2]+
rnorm(100,0,8))

## Find a 95 percent confidence region for the maximum of a quadratic polynomial
## fitted to these data

out<-OptRegionQuad(X=X,y=y,nosim=200,LB=c(-2,-2),UB=c(2,2), xlab="X1",ylab="X2")

## Example 2: a mixture-amount experiment in two components (Drug dataset) with
## non-normal data. Note triangular experimental region. Resulting 95%
## confidence region is pushed against the constraint and results in a
## "thin line"

out<-OptRegionQuad(X=Drug[,1:2],y=Drug[,3],nosim=500,LB=c(0,0),UB=c(0.08,11),
 xlab="Component 1 (mg.)",ylab="Component 2 (mg.)",triangularRegion = TRUE,
 vertex1 = c(0.02,11),vertex2 = c(0.08,1.8),outputPDFFile="Mixture_plot.pdf")
```

`OptRegionTps` *Computes Confidence Regions of Optima of Thin Plate Spline Models*

Description

Computes and displays the confidence region on the location of the optima of a thin plate spline surface in 2 factors using bootstrapping.

Usage

```
OptRegionTps(X, y, lambda=0.04, nosim=1000, alpha=0.05, LB, UB,
triangularRegion=FALSE, vertex1=NULL, vertex2=NULL, maximization=TRUE, xlab=
"Protein eaten, mg",ylab="Carbohydrate eaten, mg",outputPDFFile="CRplot.pdf",
outputOptimaFile="Optima.txt")
```

Arguments

X	nx2 matrix with the values of the 2 regressors (experimental factors) in the n observations
y	nx1 vector of response value observations
lambda	penalization parameter (larger values implies more smoothing). Default is 0.04
nosim	number of simulations (default=1000)
alpha	confidence level (0<alpha<1; default=0.05)
LB,UB	2x1 vectors of lower and upper bounds for search region where optima may lie
triangularRegion	logical: if TRUE it will constrain the optimum points to lie inside a triangle defined by the coordinates (0,0), and those in "vertex1", and "vertex2", see below (in addition to being constrained to lie inside the region defined by LB and UB). NOTE: use TRUE when the treatments form a triangular experimental region in shape. If FALSE, maxima will only be constrained to lie inside the rectangular region defined by LB and UB. Default is FALSE.
vertex1,vertex2	2x1 vectors with coordinates defining two of the 3 vertices of a triangular region. Must be provided if triangularRegion is TRUE (NOTE: vertices numbered clockwise, vertex0=c(0,0) always)
maximization	logical: if TRUE (default) it maximizes it FALSE it minimizes
xlab	text label for x axis in confidence region plot (default: "Protein eaten, mg")
ylab	text label for y axis in confidence region plot (default: "Carbohydrates eaten, mg")
outputPDFFile	name of the PDF file where the CR plot is saved (default: "CRplot.pdf")
outputOptimaFile	name of the text file containing the coordinates of all the optima found (same information as in output vector xin, see below)

Details

This program approximates the confidence region (CR) of the location of the optimum of a Thin Plate Spline (TPS) in 2 regressors x constrained inside a rectangular region defined by LB and UB. If triangularRegion=TRUE it will also contain the optimum to lie inside the experimental region assumed to be well approximated by a triangle. The CR is generated point wise by bootstrapping the residuals of a TPS fit to the given (X,y) data, refitting Tps models, and solving the corresponding constrained maximization (or minimization) problems. The confidence region is approximated by the convex hull of all the solutions (x1*,x2*) found.

Usage assuming all default options:

out<-OptRegionTps(X=X,y=y,LB=LB,UB=UB)

Value

meanPoint	a 2x1 vector with the coordinates of the mean optimum point
xin	an mx2 matrix with the x,y coordinates of all simulated #points that belong to the confidence region (dim(m) is (1-alpha)*nosim)

Note

Upon completion, a PDF file containing the CR plot with name as set in ouputPDFFile is created. A text file named as set in outputOptimaFile with all xin values is created too.

Author(s)

Enrique del Castillo <exd13@psu.edu>, John Hunt <J.Hunt@exeter.ac.uk> and James Rapkin <jr297@exeter.ac.uk>

References

Del Castillo, E., Hunt, J., Rapkin, J., and Zarmehri, S. , "Confidence regions for the location of response surface optima: the R package OptimaRegion".

Examples

```
## Not run:
## Example 1: randomly generated 2-variable response surface data

X<-cbind(runif(100,-2,2),runif(100,-2,2))

y<-as.matrix(72-11.78*X[,1]+0.74*X[,2]-7.25*X[,1]^2-7.55*X[,2]^2-4.85*X[,1]*X[,2]+
rnorm(100,0,8))

## Find a 95 percent confidence region for the maximum of a Thin Plate Spline
model fitted to these data

out<-OptRegionTps(X=X,y=y,nosim=200,LB=c(-2,-2),UB=c(2,2), xlab="X1",ylab="X2")

## Example 2: a mixture-amount experiment in two components (Drug dataset) with
non-normal data. Note triangular experimental region. Resulting 95p confidence
region of the maxima of a TPS model has area > 0. Contrast with region for
quadratic polynomial model. Note: 500 bootstrap iterations may take a few minutes.

out<-OptRegionTps(X=Drug[,1:2],y=Drug[,3],nosim=500,lambda=0.05,LB=c(0,0),UB=
 c(0.08,11), xlab="Component 1 (mg.)",ylab="Component 2 (mg.)",triangularRegion
 = TRUE,vertex1 = c(0.02,11),vertex2 = c(0.08,1.8),outputPDFFile=
 "Mixture_plot.pdf")

## End(Not run)
```

---

plotConvexHull *Computes and displays the convex hull of a set of 2-dimensional points*

Description

Given a vector of 2-dimensional coordinates, computes and displays the convex hull formed by these points.

Usage

```
plotConvexHull(xin, LB, UB, xlab, ylab)
```

Arguments

xin	n x 2 vector of coordinate points
LB	2x1 vector of lower bounds for the x,y region where the convex hull is to be plot. Required.
UB	2x1 vector of lower bounds for the x,y region where the convex hull is to be plot. Required.
xlab	Label for x axis.
ylab	Label for y axis.

Value

An integer vector giving the indices of the unique points lying on the convex hull, in clockwise order. (The first will be returned for duplicate points.). Same as returned by chull(xin).

Note

Function also plots the convex hull upon return.

Author(s)

E. del Castillo

Examples

```
## Not run:
## Generate some random 2-dimensional point set

X<-cbind(runif(100,-2,2),runif(100,-2,2))

## Compute and plot convex hull

plotConvexHull(xin=X, LB=c(-4,-4), UB=c(4,4), xlab='X', ylab='Y')

## End(Not run)
```

Index