Package: DoubleCone (via r-universe)

August 30, 2024

Type Package

Title Test Against Parametric Regression Function

Version 1.1

Date 2017-10-02

Author Mary C Meyer, Bodhisattva Sen

Maintainer Mary C Meyer <meyer@stat.colostate.edu>

Description Performs hypothesis tests concerning a regression function in a least-squares model, where the null is a parametric function, and the alternative is the union of large-dimensional convex polyhedral cones. See Bodhisattva Sen and Mary C Meyer (2016) <doi:10.1111/rssb.12178> for more details.

Depends graphics, grDevices, stats, utils, coneproj (>= 1.12), Matrix, MASS

NeedsCompilation no

Repository CRAN

Date/Publication 2017-10-02 18:52:54 UTC

Contents

DoubleCone-package — *Test against a Parametric Function*

Description

Given a response and predictors, the null hypothesis of a parametric regression function is tested versus a large-dimensional alternative in the form of a union of polyhedral convex cones.

Details

Package:	DoubleCone
Type:	Package
Version:	1.0
Date:	2013-10-24
License:	GPL-2 | GPL-3

The `doubconetest` function is the generic version. The user provides an irreducible constraint matrix that defines two convex cones; the intersection of the cones is the null space of the matrix. The function provides a p-value for the test that the expected value of a vector is in the null space using the double-cone alternative.

Given a vector `y` and a design matrix `X`, the `agconst` function performs a test of the null hypothesis that the expected value of `y` is constant versus the alternative that it is monotone (increasing or decreasing) in each of the predictors.

The function `partlintest` performs a test of a linear model versus a partial linear model, using a double-cone alternative.

Author(s)

Mary C Meyer and Bodhisattva Sen Maintainer: Mary C Meyer <meyer@stat.colostate.edu>

References

TBA

adhd — *Sub-clinical ADHD behaviors and classroom functioning in school-age children*

Description

Observations on children aged 9-11 in classroom settings, for a study on the effects of sub-clinical hyperactive and inattentive behaviors on social and academic functioning.

Usage

```
data(adhd)
```

Format

A data frame with 686 observations on the following 4 variables.

`sex` 1=boy; 2=girl

`ethn` 1=Colombian, 2=African American, 3=Hispanic American, 5=European American

`hypb` Classroom hyperactive behavior level

`fcn` A measure of social and academic functioning

Source

Brewis, A.A. Schmidt, K.L., and Meyer, M.C. (2000) ADHD-type behavior and harmful dysfunction in childhood: a cross-cultural model, American Anthropologist, 102(4), pp823-828.

Examples

```
data(adhd)
plot(adhd$hypb,adhd$fcn)
```

`agconst` *Test null hypothesis of constant regression function against a general, high-dimensional alternative*

Description

Given a response and 1-3 predictors, the function will test the null hypothesis that the response and predictors are not related (i.e., regression function is constant), against the alternative that the regression function is monotone in each of the predictors. For one predictor, the alternative set is a double cone; for two predictors the alternative set is a quadruple cone, and an octuple cone alternative is used when there are three predictors.

Usage

```
agconst(y, xmat, nsim = 1000)
```

Arguments

`y`	A numeric response vector, length n
`xmat`	an n by k design matrix, full column rank, where k=1,2, or 3.
`nsim`	The number of data sets simulated under the null hypothesis, to estimate the null distribution of the test statistic. The default is 1000, make this larger if a more precise p-value is desired.

Details

For one predictor, the set of non-decreasing regression functions can be described by an n-dimensional convex polyhedral cone, and the set of non-increasing regression functions is the "opposite" cone. The one-dimensional null space is the intersection of these cones. For two predictors, the alternative set consists of four cones, defined by combinations of increasing/decreasing assumptions, and for three predictors we have eight cones.

Value

`pval`	The p-value for the test: H0: constant regression function
`p1` through `p8`	monotone fits – only p1 and p2 are returned for one predictor, etc.
`thetahat`	The least-squares alternative fit – i.e., the projection onto the multiple-cone alternative

Author(s)

Mary C Meyer and Bodhisattva Sen

References

TBA

See Also

`doubconetest`,`partlintest`

Examples

```
 n=100
 x1=runif(n);x2=runif(n);xmat=cbind(x1,x2)
 mu=1:n;for(i in 1:n){mu[i]=20*max(x1[i]-2/3,x2[i]-2/3,0)^2}
 x1g=1:21/22;x2g=x1g
 par(mar=c(1,1,1,1))
 y=mu+rnorm(n)
 ans=agconst(y,xmat,nsim=0)
 grfit=matrix(nrow=21,ncol=21)
 for(i in 1:21){for(j in 1:21){
   if(sum(x1>=x1g[i]&x2>=x2g[j])>0){
    if(sum(x1<=x1g[i]&x2<=x2g[j])>0){
     f1=min(ans$thetahat[x1>=x1g[i]&x2>=x2g[j]])
     f2=max(ans$thetahat[x1<=x1g[i]&x2<=x2g[j]])
     grfit[i,j]=(f1+f2)/2
    }else{
     grfit[i,j]=min(ans$thetahat)
    }
   }else{grfit[i,j]=max(ans$thetahat)}
 }}
 persp(x1g,x2g,grfit,th=-50,tick="detailed",xlab="x1",ylab="x2",zlab="mu")
##to get p-value for test against constant function:
# ans=agconst(y,xmat,nsim=1000)
# ans$pval
```

derby — *Kentucky Derby Winner Speed*

Description

The Speeds of the Winning Horses in the Kentucky Derby, 1896-2012

Usage

```
data(derby)
```

Format

A data frame with 117 observations on the following 4 variables.

- `speed` winning speed
- `year` year of race
- `cond` track condition with levels `fast` `good` `heav` `mudd` `slop` `slow`
- `name` Name of the winning horse

Source

www.kentuckyderby.com

Examples

```
data(derby)
n=length(derby$year)
track=1:n*0+1
track[derby$cond=="good"]=2
track[derby$cond=="fast"]=3
plot(derby$year,derby$speed,col=track)
```

doubconetest — *Test for a vector being in the null space of a double cone*

Description

Given an n-vector y and the model y=m+e, and an m by n "irreducible" matrix amat, test the null hypothesis that the vector m is in the null space of amat.

Usage

```
doubconetest(y, amat, nsim = 1000)
```

Arguments

`y`	a vector of length n
`amat`	an m by n "irreducible" matrix
`nsim`	number of simulations to approximate null distribution – default is 1000, but choose more if a more precise p-value is desired

Details

The matrix amat defines a polyhedral convex cone of vectors x such that amat%*%x>=0, and also the opposite cone amat%*%x<=0. The linear space C is those x such that amat%*%x=0. The function provides a p-value for the null hypothesis that m=E(y) is in C, versus the alternative that it is in one of the two cones defined by amat.

Value

`pval`	The p-value for the test
`p0`	The least-squares fit under the null hypothesis
`p1`	The least-squares fit to the "positive" cone
`p2`	The least-squares fit to the "negative" cone

Author(s)

Mary C Meyer and Bodhisattva Sen

References

TBA, Meyer, M.C. (1999) An Extension of the Mixed Primal-Dual Bases Algorithm to the Case of More Constraints than Dimensions, Journal of Statistical Planning and Inference, 81, pp13-31.

See Also

agconst,partlintest

Examples

```
## test against a constant function
n=100
x=1:n/n
mu=4-5*(x-1/2)^2
y=mu+rnorm(n)
amat=matrix(0,nrow=n-1,ncol=n)
for(i in 1:(n-1)){amat[i,i]=-1;amat[i,i+1]=1}
ans=doubconetest(y,amat)
ans$pval
plot(x,y,col="slategray");lines(x,mu,lty=3,col=3)
lines(x,ans$p1,col=2)
lines(x,ans$p2,col=4)
```

partlintest *Tests linear versus partial linear model*

Description

Given a response y, a predictor x, and covariates z, the model y=m(x) +b'z +e is considered, where e is a mean-zero random error. There are three options for the null hypothesis: h0=0 tests m(x) is constant; h0=1 tests m(x) is linear, and h0=2 tests m(x) is quadratic. The (respective) alternatives are: m(x) is increasing or decreasing, m(x) is convex or concave, and m(x) is hyper-convex or hyper-concave (referring to the third derivative of m).

Usage

```
partlintest(x, y, zmat, h0 = 0, nsim = 1000)
```

Arguments

x	a vector of length n; this is the main predictor of interest
y	a vector of length n; this is the response
zmat	an n by k matrix of covariates, should be full column rank .
h0	An indicator of what null hypothesis is to be tested: h0=0 for the null hypothesis: m(x) is constant; h0=1 tests m(x) is linear, and h0=2 tests m(x) is quadratic.
nsim	The number of simulations used in creating the null distribution of the test statistic. The default is nsim=1000, if a more precise p-value is desired, make nsim larger.

Details

For the constant null hypothesis, the alternative fit is either the monotone increasing or monotone decreasing fit – whichever minimizes the sum of squared residuals. For the linear null hypothsis, the alternative fit is either convex or concave, and for the quadratic null hypothesis, the alternative fit is constrained so that the third derivative is either positive or negative over the range of x-values.

Value

pval	The p-value for the test
p0	The null hypothesis fit
p1	The "positive" fit
p2	The "negative" fit

Author(s)

Mary C Meyer and Bodhisattva Sen

References

TBA

See Also

agconst,doubconetest

Examples

```
data(derby)
n=length(derby$speed)
zmat=matrix(0,nrow=n,ncol=2);zvec=1:n*0+1
zmat[derby$cond=="good",1]=1;zvec[derby$cond=="good"]=2
zmat[derby$cond=="fast",2]=1;zvec[derby$cond=="fast"]=3
ans=partlintest(derby$year,derby$speed,zmat,h0=2)
ans$pval
par(mar=c(4,4,1,1));par(mfrow=c(1,2))
plot(derby$year,derby$speed,col=zvec,pch=zvec)
points(derby$year,ans$p0,pch=20,col=zvec)
title("Null fit")
legend(1980,51.6,pch=3:1,col=3:1,legend=c("fast","good","slow"))
plot(derby$year,derby$speed,col=zvec,pch=zvec)
points(derby$year,ans$p1,pch=20,col=zvec)
title("Alternative fit")

data(adhd)
n=length(adhd$sex)
zmat=matrix(0,nrow=n,ncol=2)
zmat[adhd$sex==1,1]=1
zmat[adhd$ethn<5,2]=1
ans=partlintest(adhd$hypb,adhd$fcn,zmat,h0=1)
ans$pval
cols=c("pink3","lightskyblue3")
plot(adhd$hypb,adhd$fcn,col=cols[zmat[,1]+1],pch=zmat[,2]+1,
xlab="Hyperactive behavior level",ylab="Social and Academic Function Score")
cols2=c(2,4)
points(adhd$hypb,ans$p1,col=cols2[zmat[,1]+1],pch=20)
```

Index