Package 'NPsimex'

February 19, 2015

Type Package

Title Nonparametric Smoothing for contaminated data using Simulation-Extrapolation

Version 0.2-1

Date 2011-11-3

Author Xiao-Feng Wang

Maintainer Xiao-Feng Wang <wangx6@ccf.org>

Description This package contains a collection of functions to to perform nonparametric deconvolution using simulation extrapolation (SIMEX). We propose an estimator that adopts the SIMEX idea but bypasses the simulation step in the original SIMEX algorithm. There is no bandwidth parameter and the estimate is determined by appropriately selecting ``design points". See details in: Wang, X.F., Sun, J. and Fan, Z. (2011). Deconvolution density estimation with heteroscedastic errors using SIMEX.

License GPL (>= 3)

Repository CRAN

Date/Publication 2011-11-04 20:30:22

NeedsCompilation yes

R topics documented:

lambda.select *A rule of thumb lambda_1 selection in density estimation with SIMEX*

Description

To compute the lambda_1 using the rule-of-thumb methods.

Usage

```
lambda.select(W,msigma, method="SJ", na.rm = FALSE, ...)
```

Arguments

W	The observed data. It is a vector of length at least 10.
msigma	The standard deviation σ of measurement error. It is a single positive numeric value.
method	Specifies the bandwidth of 'W'. It can be a single numeric value which has been pre-determined; or computed with the specific density bandwidth selector: 'nrd0', 'nrd', 'ucv', 'bcv', 'SJ'.
na.rm	is set to FALSE by default: no NA value is allowed.
...	control

Author(s)

X.F. Wang <wangx6@ccf.org>

References

Wang, X.F., Sun, J. and Fan, Z. (2011). Deconvolution density estimation with heteroscedastic errors using SIMEX.

See Also

simex.density, simex.H.density.

Examples

```
N <- 1000
set.seed(123); X <- c(rnorm(N/2, mean=-2), rnorm(N/2,mean=2)); U <- rnorm(N,sd=1)
msigma <- 0.5
#msigma <- runif(N,min=0.3,max=0.5)
W <- X + msigma*U
lambda1 <- lambda.select(W, msigma=msigma)
```

plot.NPsimex *Plot a NPsimex Object*

Description

To plot a "NPsimex" object generated by simex.density(...), or simexH.density(...).

Details

This function is to plot the estimated function generated by simex.density(...), or simexH.density(...)

Author(s)

X.F. Wang <wangx6@ccf.org>

See Also

simex.density, simex.H.density.

print.NPsimex *Print a NPsimex Object*

Description

To print a "NPsimex" object generated by simex.density(...), or simexH.density(...).

Details

This function is to print the summary description from the object generated by simex.density(...), or simexH.density(...)

Author(s)

X.F. Wang <wangx6@ccf.org>

See Also

simex.density, simex.H.density.

simex.density *Estimating probability density function from data with homoscedastic measurement error using SIMEX*

Description

To compute the probability density function from data contaminated with homoscedastic measurement error using SIMEX.

Usage

```
simex.density(W, msigma, x, from, to, n.user=128, n.lambda=50,
lambda="SJ", span=8, adjust=1, na.rm = FALSE, ...)
```

Arguments

W	The observed data. It is a vector of length at least 10.
msigma	The standard deviation σ of measurement error. It is a single positive numeric value.
x	x is user-defined grids where the PDF will be evaluated.
from	the starting point where the PDF is to be evaluated.
to	the starting point where the PDF is to be evaluated.
n.user	number of points where the PDF is to be evaluated.
n.lambda	number of points of lambda's.
lambda	Specifies the first lambda. It can be a single numeric value which has been predetermined; or computed with the specific density bandwidth selector: 'nrd0', 'nrd', 'ucv', 'bcv', 'SJ'.
span	Specifies the span of lambda.
adjust	adjust the range there the PDF is to be evaluated. By default, $adjust = 1$.
na.rm	is set to FALSE by default: no NA value is allowed.
...	control

Value

An object of class "NPsimex".

Author(s)

X.F. Wang <wangx6@ccf.org>

References

Wang, X.F., Sun, J. and Fan, Z. (2011). Deconvolution density estimation with heteroscedastic errors using SIMEX.

See Also

simex.density, lambda.select, span.select.

Examples

```
############### Homoscadestic error
N <- 1000
set.seed(123); X <- c(rnorm(N/2, mean=-2), rnorm(N/2,mean=2)); U <- rnorm(N,sd=1)
msigma <- 0.5
W <- X + msigma*U

plot.simex.density <- function(X.simex,X,...){
plot(X.simex$x, X.simex$y, type="l", xlab="x", ylab="density", lwd=3, lty=2, col=2,...)
lines(density(X, bw="SJ"), lwd=3)
}

par(mfrow=c(2,2))
X.simex1 <- simex.density(W, msigma=msigma, adjust=1, n.lambda=50, span=1)
plot.simex.density(X.simex1, X, ylim=c(0,0.25))

X.simex2 <- simex.density(W, msigma=msigma, adjust=1, n.lambda=50, span=3)
plot.simex.density(X.simex2, X, ylim=c(0,0.25))

X.simex3 <- simex.density(W, msigma=msigma, adjust=1, n.lambda=50, span=8)
plot.simex.density(X.simex3, X, ylim=c(0,0.25))

X.simex4 <- simex.density(W, msigma=msigma, adjust=1, n.lambda=50, span=35)
plot.simex.density(X.simex4, X, ylim=c(0,0.25))
```

---

simex.H.density — *Estimating probability density function from data with heteroscedastic measurement error using SIMEX*

Description

To compute the probability density function from data contaminated with heteroscedastic measurement error using SIMEX.

Usage

```
simex.H.density(W, msigma, x, from, to, n.user=128, n.lambda=50,
lambda="SJ", span=8, adjust=1, na.rm = FALSE, ...)
```

Arguments

W	The observed data. It is a vector of length at least 10.
msigma	The standard deviation σ of measurement error. It is a vector of standard deviations having the same length as W.
x	x is user-defined grids where the PDF will be evaluated.
from	the starting point where the PDF is to be evaluated.
to	the starting point where the PDF is to be evaluated.
n.user	number of points where the PDF is to be evaluated.
n.lambda	number of points of lambda's.
lambda	Specifies the first lambda. It can be a single numeric value which has been predetermined; or computed with the specific density bandwidth selector: 'nrd0', 'nrd', 'ucv', 'bcv', 'SJ'.
span	Specifies the span of lambda.
adjust	adjust the range there the PDF is to be evaluated. By default, $adjust = 1$.
na.rm	is set to FALSE by default: no NA value is allowed.
...	control

Value

An object of class "NPsimex".

Author(s)

X.F. Wang <wangx6@ccf.org>

References

Wang, X.F., Sun, J. and Fan, Z.(2011). Deconvolution density estimation with heteroscedastic errors using SIMEX.

See Also

simex.density, lambda.select, span.H.select.

Examples

```
################ Heteroscadestic error
N <- 1000
set.seed(123); X <- c(rnorm(N/2, mean=-2), rnorm(N/2,mean=2)); U <- rnorm(N,sd=1); msigma <- runif(N,min=0.3,max=
W <- X + msigma*U

plot.simex.density <- function(X.simex,X,...){
plot(X.simex$x, X.simex$y, type="l", xlab="x", ylab="density", lwd=3, lty=2, col=2,...)
lines(density(X, bw="SJ"), lwd=3)
}
```

```
par(mfrow=c(2,2))
X.simex1 <- simex.H.density(W, msigma=msigma, adjust=1, n.lambda=50, span=1)
plot.simex.density(X.simex1, X, ylim=c(0,0.25))

X.simex2 <- simex.H.density(W, msigma=msigma, adjust=1, n.lambda=50, span=3)
plot.simex.density(X.simex2, X, ylim=c(0,0.25))

X.simex3 <- simex.H.density(W, msigma=msigma, adjust=1, n.lambda=50, span=8)
plot.simex.density(X.simex3, X, ylim=c(0,0.25))

X.simex4 <- simex.H.density(W, msigma=msigma, adjust=1, n.lambda=50, span=35)
plot.simex.density(X.simex4, X, ylim=c(0,0.25))
```

`span.H.select` *A method to select the span of the sequence of lambda's for the case of heteroscedastic error*

Description

To compute the optimal span of the sequence of lambda's for the case of heteroscedastic error.

Usage

```
span.H.select(W, msigma, span=c(2,4,6,8,10,12,25), approx=FALSE, from=min(W), to=max(W), n.user=128,
```

Arguments

`W`	The observed data. It is a vector of length at least 10.
`msigma`	The standard deviation σ of measurement error. It is a single positive numeric value.
`span`	span is user-defined grids where the PDF will be evaluated.
`approx`	FALSE by default; if TRUE, the lose function uses the same one as the one of the homoscedastic error case with the average of measurement error variances
`from`	the starting point where the PDF is to be evaluated.
`to`	the starting point where the PDF is to be evaluated.
`n.user`	number of points where the PDF is to be evaluated.
`n.lambda`	number of points of lambda's.
`lambda`	Specifies the first lambda. It can be a single numeric value which has been pre-determined; or computed with the specific density bandwidth selector: 'nrd0', 'nrd', 'ucv', 'bcv', 'SJ'.
`bw`	Specifies the bandwidth of 'W'. It can be a single numeric value which has been pre-determined; or computed with the specific density bandwidth selector: 'nrd0', 'nrd', 'ucv', 'bcv', 'SJ'.
`adjust`	adjust the range there the PDF is to be evaluated. By default, $adjust = 1$.
`na.rm`	is set to FALSE by default: no NA value is allowed.
`...`	control

Author(s)

X.F. Wang <wangx6@ccf.org>

References

Wang, X.F., Sun, J. and Fan, Z. (2011). Deconvolution density estimation with heteroscedastic errors using SIMEX.

See Also

simex.density.

Examples

```
plot.simex.density <- function(X.simex,X,...){
plot(X.simex$x, X.simex$y, type="l", xlab="x", ylab="density", lwd=3, lty=2, col=2,...)
lines(density(X, bw="SJ"), lwd=3)
}

################ Heteroscadestic error
N <- 1000
set.seed(123); X <- c(rnorm(N/2, mean=-2), rnorm(N/2,mean=2)); U <- rnorm(N,sd=1); msigma <- runif(N,min=0.3,max=
W <- X + msigma*U

plot.simex.density <- function(X.simex,X,...){
plot(X.simex$x, X.simex$y, type="l", xlab="x", ylab="density", lwd=3, lty=2, col=2,...)
lines(density(X, bw="SJ"), lwd=3)
}


#---- Select the optimal lambda span
par(mfrow=c(1,2))
spans <- span.H.select(W, msigma, span=c(2,4,6,8,10,12,16,25), approx=TRUE)
plot(spans$span, spans$ISE, type="o", xlab="span", ylab="ISE")

X.simex <- simex.H.density(W, msigma=msigma, adjust=1, n.lambda=50, span=spans$span[order(spans$ISE)[1]])
plot.simex.density(X.simex, X,ylim=c(0,0.25))

## more computational time needed if approx=FALSE.

# spans <- span.H.select(W, msigma, span=c(2,4,6,8,10,12,16,25), approx=FALSE)
# plot(spans$span, spans$ISE, type="o", xlab="span", ylab="ISE")
#
# X.simex <- simex.H.density(W, msigma=msigma, adjust=1, n.lambda=50, span=spans$span[order(spans$ISE)[1]])
# plot.simex.density(X.simex, X,ylim=c(0,0.25))
```

span.select *A method to select the span of the sequence of lambda's for the case of homoscedastic error*

Description

To compute the optimal span of the sequence of lambda's for the case of homoscedastic error.

Usage

```
span.select(W, msigma, span=c(2,4,6,8,10,12,25), from=min(W), to=max(W), n.user=128, n.lambda=50, la
```

Arguments

W	The observed data. It is a vector of length at least 10.
msigma	The standard deviation σ of measurement error. It is a single positive numeric value.
span	span is a vector of user-defined grids for spans.
from	the starting point where the PDF is to be evaluated.
to	the starting point where the PDF is to be evaluated.
n.user	number of points where the PDF is to be evaluated.
n.lambda	number of points of lambda's.
lambda	Specifies the first lambda. It can be a single numeric value which has been pre-determined; or computed with the specific density bandwidth selector: 'nrd0', 'nrd', 'ucv', 'bcv', 'SJ'.
bw	Specifies the bandwidth of 'W'. It can be a single numeric value which has been pre-determined; or computed with the specific density bandwidth selector: 'nrd0', 'nrd', 'ucv', 'bcv', 'SJ'.
adjust	adjust the range there the PDF is to be evaluated. By default, $adjust = 1$.
na.rm	is set to FALSE by default: no NA value is allowed.
...	control

Author(s)

X.F. Wang <wangx6@ccf.org>

References

Wang, X.F., Sun, J. and Fan, Z. (2011). Deconvolution density estimation with heteroscedastic errors using SIMEX.

See Also

simex.density.

Examples

```
############### Homoscadestic error
N <- 1000
set.seed(123); X <- c(rnorm(N/2, mean=-2), rnorm(N/2,mean=2)); U <- rnorm(N,sd=1)
msigma <- 0.5
W <- X + msigma*U

plot.simex.density <- function(X.simex,X,...){
plot(X.simex$x, X.simex$y, type="l", xlab="x", ylab="density", lwd=3, lty=2, col=2,...)
lines(density(X, bw="SJ"), lwd=3)
}

#---- Select the optimal lambda span
par(mfrow=c(1,2))
spans <- span.select(W, msigma)
plot(spans$span, spans$ISE, type="o", xlab="span", ylab="ISE")

X.simex <- simex.density(W, msigma=msigma, adjust=1, n.lambda=50, span=spans$span[order(spans$ISE)[1]])
plot.simex.density(X.simex, X, ylim=c(0,0.25))
```

Index