Package ‘monreg’

October 13, 2022

Version 0.1.4
Date 2020-04-25
Title Nonparametric Monotone Regression
Author This version Kay Pilz and Steffanie Titoff. Earlier developments by Holger Dette and Kay Pilz.
Maintainer Scott Kostyshak <scott.kostyshak@gmail.com>
Depends R (>= 2.0.0)
Encoding latin1
License GPL (>= 2)
NeedsCompilation yes
URL https://gitlab.com/scottkosty/monreg
BugReports https://gitlab.com/scottkosty/monreg/issues
Repository CRAN
Date/Publication 2020-04-26 07:00:09 UTC

R topics documented:

monreg ................................................................. 2
monvardiff ............................................................ 4
monvarresid ......................................................... 6

Index 9
monreg provides a strictly monotone estimator of the regression function based on the nonparametric regression model.

Usage

```r
monreg(x, y, a = min(x), b = max(x), N = length(x), t = length(x),
       hd, Kd = "epanech", hr, Kr = "epanech", degree = 1,
       inverse = 0, monotonie = "isoton")
```

Arguments

- `x`: vector containing the x-values (design points) of a sample
- `y`: vector containing the y-values (response) of a sample
- `a`: lower bound of the support of the design points density function, or smallest fixed design point
- `b`: upper bound of the support of the design points density function, or largest fixed design point
- `N`: number or vector of evaluation points of the unconstrained nonparametric regression estimator (e.g. Nadaraya-Watson estimator)
- `t`: number or vector of points where the monotone estimation is computed
- `hd`: bandwidth of kernel $K_d$ of the density estimation step
- `Kd`: Kernel for the density estimation step (monotonization step). 'epanech' for Epanechnikov, 'rectangle' for rectangle, 'biweight' for biweight, 'triweight' for triweight, 'triangle' for triangle, 'cosine' for cosine kernel
- `hr`: bandwidth of kernel $K_r$ of the regression estimation step.
- `degree`: Determines the method for the unconstrained estimation. '0' for the classical Nadaraya-Watson estimate, '1' for the local linear estimate. As well degree can be the vector of the unconditional estimator provided by the user for the design points given in the vector `N`
- `inverse`: For '0' the original regression function is estimated, for '1' the inverse of the regression function is estimated.
- `monotonie`: Determines the type of monotonicity. 'isoton' if the regression function is assumed to be isotone, 'antinton' if the regression function is assumed to be antitone.
Details

Nonparametric regression models are of the form $Y_i = m(X_i) + \sigma(X_i) \cdot \varepsilon_i$, where $m$ is the regression function and $\sigma$ the variance function. `monreg` performs a monotone estimate of the unknown regression function $m$. `monreg` first estimates $m$ by an unconstrained nonparametric method, the classical Nadaraya-Watson estimate or the local-linear estimate (unless the user decides to pass his or her own estimate). In a second step the inverse of the (monotone) regression function is calculated, by monotonizing this unconstrained estimate. With the above notation and $\hat{m}$ for the unconstrained estimate, the second step writes as follows,

$$\hat{m}^{-1}_t = \frac{1}{N h_d} \sum_{i=1}^{N} \int_{-\infty}^{t} K_d \left( \frac{\hat{m}(t/N) - u}{h_d} \right) du.$$

Finally, the monotone estimate achieved by inversion of $\hat{m}^{-1}_t$.

Value

`monreg` returns a list of values

- `xs` the input values $x$, standardized on the interval $[0, 1]$
- `y` input variable $y$
- `z` the points, for which the unconstrained function is estimated
- `t` the points, for which the monotone function values will be estimated
- `length.x` length of the vector $x$
- `length.z` length of the vector $z$
- `length.t` length of the vector $t$
- `hd` bandwidth used with the Kernel $K_d$
- `hr` bandwidth used with the Kernel $K_r$
- `Kd` kernel used for the monotonization step
- `Kr` kernel used for the initial unconstrained regression estimate
- `degree` method, which was used for the unconstrained regression estimate
- `ldeg.vektor` length of the vector degree. If `ldeg.vektor` is not equal to 1 the user provided the vector of the unconditional estimator for the design points given in the vector $N$
- `inverse` indicates, if the origin regression function or its inverse has been estimated
- `estimation` the monotone estimate at the design points $t$

...

Author(s)

This R Package was developed by Kay Pilz and Stefanie Titoff. Earlier developements of the estimator were made by Holger Dette and Kay Pilz.

See Also

`monvardiff` and `monvarresid` for monotone variance function estimation.
Examples

```r
x <- rnorm(100)
y <- x + rnorm(100)
mon1 <- monreg(x, y, hd = .5, hr = .5)
plot(mon1$t, mon1$estimation)
```

---

monvardiff

*Estimating Monotone Variance Functions Using Pseudo-Residuals*

**Description**

monvardiff provides a strictly monotone estimator of the variance function based on the nonparametric regression model.

**Usage**

```r
monvardiff(x, y, a = min(x), b = max(x), N = length(x), t = length(x), r = 2, hr, Kr = "epanech", hd, Kd = "epanech", degree = 1, inverse = 0, monotonie = "isoton")
```

**Arguments**

- `x`: vector containing the x-values (design points) of a sample
- `y`: vector containing the y-values (response) of a sample
- `a`: lower bound of the support of the design points density function, or smallest fixed design point
- `b`: upper bound of the support of the design points density function, or largest fixed design point
- `N`: number or vector of evaluation points of the unconstrained nonparametric variance estimator (e.g. Nadaraya-Watson estimator)
- `t`: number or vector of points where the monotone estimation is computed
- `r`: order of the difference scheme, i.e. weights $d_0, ..., d_r$ to calculate the pseudo-residuals
- `hr`: bandwidth of kernel $K_r$ of the variance estimation step
- `Kr`: Kernel for the variance estimation step (unconstrained estimation). 'epanech' for Epanechnikov, 'rectangle' for rectangle, 'biweight' for biweight, 'triweight' for triweight, 'triangle' for triangle, 'cosine' for cosine kernel
- `hd`: bandwidth of kernel $K_d$ of the density estimation step
- `Kd`: Kernel for the density estimation step (monotonization step). 'epanech' for Epanechnikov, 'rectangle' for rectangle, 'biweight' for biweight, 'triweight' for triweight, 'triangle' for triangle, 'cosine' for cosine kernel
- `degree`: determines the method for the unconstrained variance estimation. '0' for the classical Nadaraya-Watson estimate, '1' for the local linear estimate. As well degree can be the vector of the unconditional estimator provided by the user for the design points given in the vector `N`
inverse for '0' the original variance function is estimated, for '1' the inverse of the variance function is estimated.

monotonie determines the type of monotonicity. 'isoton' if the variance function is assumed to be isotone, 'antinton' if the variance function is assumed to be antitone.

Details

Nonparametric regression models are of the form \( Y_i = m(X_i) + \sigma(X_i) \cdot \varepsilon_i \), where \( m \) is the regression function and \( \sigma \) the variance function. \texttt{monvardiff} performs a monotone estimate of the unknown variance function \( s = \sigma^2 \). \texttt{monvardiff} first estimates \( s \) by an unconstrained nonparametric method, the classical Nadaraya-Watson estimate or the local-linear estimate (unless the user decides to pass his or her own estimate). This estimation contains the usage of the Pseudo-Residuals. In a second step the inverse of the (monotone) variance function is calculated by monotonizing the unconstrained estimate from the first step. With the above notation and \( \hat{s} \) for the unconstrained estimate, the second step writes as follows,

\[
\hat{s}^{-1}_I = \frac{1}{Nh_d} \sum_{i=1}^{N} \int_{-\infty}^{t} K_d \left( \frac{\hat{s}((N)^{-1}) - u}{h_d} \right) du.
\]

Finally, the monotone estimate is achieved by inversion of \( \hat{s}^{-1}_I \).

Value

\texttt{monvardiff} returns a list of values

- \( xs \) the input values \( x \), standardized on the interval \([0, 1]\)
- \( y \) input variable \( y \)
- \( z \) the points, for which the unconstrained function is estimated
- \( t \) the points, for which the monotone variance function will be estimated
- \( \text{length.x} \) length of the vector \( x \)
- \( \text{length.z} \) length of the vector \( z \)
- \( \text{length.t} \) length of the vector \( t \)
- \( r \) order of the difference scheme, i.e. number of weights to calculate the pseudo-residuals
- \( hr \) bandwidth used with the Kernel \( K_r \)
- \( hd \) bandwidth used with the Kernel \( K_d \)
- \( Kr \) kernel used for the unconstrained variance estimate
- \( Kd \) kernel used for the monotonization step
- \( \text{degree} \) method, which was used for the unconstrained variance estimate
- \( \text{ldeg.vektor} \) length of the vector degree. If \( \text{ldeg.vektor} \) is not equal to 1 the user provided the vector of the unconditional variance estimator for the design points given in the vector \( N \)
- \( \text{inverse} \) indicates, if the origin variance function or its inverse has been estimated
- \( \text{estimation} \) the monotone estimate at the design points \( t \)
This R Package was developed by Kay Pilz and Stefanie Titoff. Earlier developments of the estimator were made by Holger Dette and Kay Pilz.

See Also
monreg for monotone regression function estimation and monvarresid for monotone variance function estimation by nonparametric residuals.

monvarresid

Estimating Monotone Variance Functions Using Nonparametric Residuals

Description
monvarresid provides a strictly monotone estimator of the variance function based on the nonparametric regression model.

Usage
monvarresid(x,y,a=min(x),b=max(x),N=length(x),t=length(x),h,K="epanech",hd,Kd="epanech", hr,Kr="epanech",mdegree=1,sdegree=1,inverse=0,monotonie="isoton")

Arguments

x vector containing the x-values (design points) of a sample
y vector containing the y-values (response) of a sample
a lower bound of the support of the design points density function, or smallest fixed design point
b upper bound of the support of the design points density function, or largest fixed design point
N number or vector of evaluation points of the unconstrained nonparametric variance estimator (e.g. Nadaraya-Watson estimator)
t number or vector of points where the monotone estimation is computed
h bandwidth of kernel $K$ of the regression estimation step
K Kernel for the regression estimation step. 'epanech' for Epanechnikov, 'rectangle' for rectangle, 'biweight' for biweight, 'triweight' for triweight, 'triangle' for triangle, 'cosine' for cosine kernel
hd bandwidth of kernel $K_d$ of the density estimation step
Kd Kernel for the density estimation step (monotonization step). 'epanech' for "Epanechnikov, 'rectangle' for rectangle, 'biweight' for biweight, 'triweight' for triweight, 'triangle' for triangle, 'cosine' for cosine kernel
hr bandwidth of kernel $K_r$ of the variance estimation step
\texttt{monvarresid}

\begin{itemize}
  \item \texttt{Kr} \hspace{1cm} \text{Kernel for the variance estimation step (unconstrained estimation). 'epanech' for "Epanechnikov, 'rectangle' for rectangle, 'biweight' for biweight, 'triweight' for triweight, 'triangle' for triangle, 'cosine' for cosine kernel.}
  
  \item \texttt{mdegree} \hspace{1cm} \text{determines the method for the regression estimation. '0' for the classical Nadaraya-Watson estimate, '1' for the local linear estimate. As well mdegree can be the vector of the estimator provided by the user for the design points given by the vector \texttt{x}}
  
  \item \texttt{sdegree} \hspace{1cm} \text{Determines the method for the unconstrained variance estimation. '0' for the classical Nadaraya-Watson estimate, '1' for the local linear estimate. As well sdegree can be the vector of the unconditional estimator provided by the user for the design points given by the vector \texttt{N}}
  
  \item \texttt{inverse} \hspace{1cm} \text{For '0' the original variance function is estimated, for '1' the inverse of the variance function is estimated.}
  
  \item \texttt{monotonie} \hspace{1cm} \text{Determines the type of monotonicity. 'isoton' if the variance function is assumed to be isotone, 'antinton' if the variance function is assumed to be antitone.}
\end{itemize}

**Details**

Nonparametric regression models are of the form \( Y_i = m(X_i) + \sigma(X_i) \varepsilon_i \), where \( m \) is the regression function and \( \sigma \) the variance function. \texttt{monvarresid} performs a monotone estimate of the unknown variance function \( s = \sigma^2 \). \texttt{monvarresid} first estimates \( m \) by an unconstrained nonparametric method, the classical Nadaraya-Watson estimate or the local-linear estimate (unless the user decides to pass his or her own estimate). In a second step an unconstrained estimation for \( s \) is performed, again by the classical Nadaraya-Watson method or the local-linear estimate (unless the user decides to pass his or her own estimate). In a third step the inverse of the (monotone) variance function is calculated, by monotonizing the unconstrained estimate from the second step. With the above notation and \( \hat{s} \) for the unconstrained estimate, the third step writes as follows,

\[ \hat{s}^{-1}_{\hat{f}} = \frac{1}{Nh_d} \sum_{i=1}^{N} \int_{-\infty}^{t} K_d \left( \frac{\hat{s} \left( \frac{X_i}{N} \right) - u}{h_d} \right) du. \]

Finally, the monotone estimate is achieved by inversion of \( \hat{s}^{-1}_{\hat{f}} \).

**Value**

\texttt{monvarresid} returns a list of values

- \texttt{xs} \hspace{1cm} the input values \( x \), standardized on the interval \([0, 1]\)
- \texttt{y} \hspace{1cm} input variable \( y \)
- \texttt{z} \hspace{1cm} the points, for which the unconstrained function is estimated
- \texttt{t} \hspace{1cm} the points, for which the monotone variance function will be estimated
- \texttt{length.x} \hspace{1cm} length of the vector \( x \)
- \texttt{length.z} \hspace{1cm} length of the vector \( z \)
- \texttt{length.t} \hspace{1cm} length of the vector \( t \)
h bandwidth used with the Kernel $K$

hd bandwidth used with the Kernel $K_d$

hr bandwidth used with the Kernel $K_r$

K kernel used for the regression estimation step

Kd kernel used for the monotonization step

Kr kernel used for the unconstrained variance estimate

mdegree method, which was used for the unconstrained regression estimate

lmdeg length of the vector mdegree. If lmdeg is not equal to 1 the user provided the vector of the unconditional regression estimator for the design points given by the vector x

sdegree method, which was used for the unconstrained variance estimate

lsdeg length of the vector sdegree. If lsdeg is not equal to 1 the user provided the vector of the unconditional variance estimator for the design points given by the vector N

inverse indicates, if the origin variance function or its inverse has been estimated

estimation the monotone estimate for the variance function at the design points $t$

Author(s)

This R Package was developed by Kay Pilz and Stefanie Titoff. Earlier developments of the estimator were made by Holger Dette and Kay Pilz.

See Also

monreg for monotone regression function estimation and monvardiff for monotone variance function estimation by differences.
Index

* nonparametric
  - monreg, 2
  - monvardiff, 4
  - monvarresid, 6

* regression
  - monreg, 2
  - monvardiff, 4
  - monvarresid, 6

* smooth
  - monreg, 2
  - monvardiff, 4
  - monvarresid, 6

* monvardiff, 4

monreg, 2
monvardiff, 4
monvarresid, 6