

Package ‘glober’

June 7, 2023

Type Package

Title Estimating Functions with Multivariate B-Splines

Version 1.0

Date 2023-05-16

Author M. E. Savino

Maintainer Mary E. Savino <mary.savino@outlook.fr>

Description Generalized Lasso applied to knot selection in multivariate B-spline Regression (GLOBER) implements a novel approach for estimating functions in a multivariate nonparametric regression model based on an adaptive knot selection for B-splines using the Generalized Lasso. For further details we refer the reader to the paper Savino, M. E. and Lévy-Leduc, C. (2023), <[arXiv:2306.00686](#)>.

License GPL-2

Encoding UTF-8

LazyData true

Depends R (>= 3.5.0), Matrix, genlasso, fda, parallel

Imports ggplot2, plot3D

Suggests knitr, markdown

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2023-06-07 13:20:02 UTC

R topics documented:

glober-package	2
f_1D	2
f_2D	3
glober.1d	3
glober.2d	4
xpred_1D	5

xpred_2D	6
x_1D	6
x_2D	6
y_1D	7
y_2D	7

Index	8
--------------	----------

glober-package	<i>Estimating Functions with Multivariate B-Splines</i>
----------------	---

Description

glober consists of two functions: "glober.1d.R" and "glober.2d.R". For further information on how to use these functions, we refer the reader to the vignette of the package.

Details

Eight datasets are also provided within this package and used in examples of this manual and in the vignette.

Author(s)

Mary E. Savino

Maintainer: Mary E. Savino <mary.savino@outlook.fr>

References

Savino, M. E. and Levy-Leduc, C. A novel approach for estimating functions in the multivariate setting based on an adaptive knot selection for B-splines with an application to a chemical system used in geoscience (2023), arXiv:2306.00686.

f_1D	<i>Output values of the evaluation of a function with one input variable and defined as a linear combination of B-splines</i>
------	---

Description

The evaluations were obtained by using the fda package with 201 input values contained in the xpred_1D.rda file.

Usage

```
data("f_1D")
```

Format

Numeric vector of 201 values.

f_2D	<i>Output values of the evaluation of a function with two input variables and defined as a linear combination of a tensor product of univariate B-splines</i>
------	---

Description

The evaluations were obtained by using the fda package with 10000 input values contained in the xpred_2D.rda file.

Usage

```
data("f_2D")
```

Format

Numeric vector of 10000 values.

glober.1d	<i>Estimation of functions with one input variable</i>
-----------	--

Description

This function implements the method described in Savino, M and Levy-Leduc, C (2023) for estimating functions in the one-dimensional setting with observations which are assumed to satisfy a nonparametric regression model. The observation points belong to a compact set of \mathbb{R} .

Usage

```
glober.1d(x, y, xpred, ord, parallel = FALSE, nb.Cores = 1)
```

Arguments

x	matrix of one column or vector containing the values of the input variables.
y	vector containing the corresponding response variable associated to the input values x.
xpred	matrix of one column or vector containing the input variables for which f has to be estimated.
ord	order of the B-spline basis used in the regression model. Default is 3 (quadratic B-splines).
parallel	logical, if TRUE then a parallelized version of the code is used. Default is FALSE.
nb.Cores	numerical, number of cores used for parallelization, if parallel is set to TRUE.

Value

festimated	estimation of f at x_{pred} .
knotSelec	selected knots used in the definition of the B-splines.
rss	residual sum-of-squares (RSS) of the model.
rsq	R-squared of the model, calculated as $1 - \frac{RSS}{TSS}$ where TSS is the total sum-of-squares of the model.

Examples

```
# --- Loading values of x --- #
data('x_1D')
# --- Loading values of the corresponding y --- #
data('y_1D')
# --- Loading values of xpred --- #
data('xpred_1D')

# --- Estimation of f at xpred --- #
glober.1d(x = x_1D, y = y_1D, xpred = xpred_1D, ord = 3, parallel = FALSE)

# --- Parallel computing --- #
glober.1d(x = x_1D, y = y_1D, xpred = xpred_1D, ord = 3, parallel = TRUE, nb.Cores = 2)
```

glober.2d

Estimation of functions with two input variables

Description

This function implements the method described in Savino, M and Levy-Leduc, C (2023) for estimating functions in the two-dimensional setting with observations which are assumed to satisfy a nonparametric regression model. The observation points belong to a compact set of \mathbb{R}^2 .

Usage

```
glober.2d(x, y, xpred, ord, parallel = FALSE, nb.Cores = 1)
```

Arguments

x	matrix of two columns containing the values of the input variables.
y	vector containing the corresponding response variable associated to the input values x.
xpred	matrix of one column or vector containing the input variables for which f has to be estimated.
ord	order of the B-spline basis used in the regression model. Default is 3 (quadratic B-splines).

parallel	logical, if TRUE then a parallelized version of the code is used. Default is FALSE.
nb.Cores	numerical, number of cores used for parallelization, if parallel is set to TRUE.

Value

festimated	estimation of f at xpred.
knotSelec	list of selected knots for each dimension used in the definition of the B-splines.
rss	residual sum-of-squares (RSS) of the model.
rsq	R-squared of the model, calculated as $1 - \frac{RSS}{TSS}$ where TSS is the total sum-of-squares of the model.

Examples

```
# --- Loading values of x --- #
data('x_2D')
# --- Loading values of the corresponding y --- #
data('y_2D')
# --- Loading values of xpred --- #
data('xpred_2D')

# --- Estimation of f at xpred --- #
glober.2d(x = x_2D, y = y_2D, xpred = xpred_2D, ord = 3, parallel = FALSE)

# --- Parallel computing --- #
glober.2d(x = x_2D, y = y_2D, xpred = xpred_2D, ord = 3, parallel = TRUE, nb.Cores = 2)
```

xpred_1D	<i>Values of the single input variable for which a function has to be estimated</i>
----------	---

Description

An example of input values containing the observation set and values for which an estimation of a given function f of one input variable is needed.

Usage

```
data("xpred_1D")
```

Format

Numeric vector of 201 values.

xpred_2D	<i>Values of the two input variables for which a function has to be estimated</i>
----------	---

Description

An example of input values containing the observation set and values for which an estimation of a given function f of two input variables is needed.

Usage

```
data("xpred_2D")
```

Format

Numeric matrix of 10000 rows and 2 columns.

x_1D	<i>Observation vector x of one variable</i>
------	--

Description

An example of observations for the estimation of a function f with one input variable.

Usage

```
data("x_1D")
```

Format

Numeric vector of 70 values.

x_2D	<i>Observation matrix x of two variables</i>
------	---

Description

An example of 100 observations for the estimation of a function f with two input variables.

Usage

```
data("x_2D")
```

Format

Numeric matrix of 100 rows and 2 columns.

y_1D	<i>Values of the response variable of the noisy observation set of one input variable</i>
------	---

Description

An example of noisy observations obtained by adding a Gaussian noise to $f(x_i)$ contained in f_1D.rda and associated to the input values contained in x_1D.rda.

Usage

```
data("y_1D")
```

Format

Numeric vector of 70 values.

y_2D	<i>Values of the response variable of the noisy observation set of two input variables</i>
------	--

Description

An example of noisy observations obtained by adding a Gaussian noise to $f(x_i)$ contained in f_2D.rda and associated to the input values contained in x_2D.rda.

Usage

```
data("y_2D")
```

Format

Numeric vector of 100 values.

Index

* package

glober-package, 2

f_1D, 2

f_2D, 3

glober (glober-package), 2

glober-package, 2

glober.1d, 3

glober.2d, 4

x_1D, 6

x_2D, 6

xpred_1D, 5

xpred_2D, 6

y_1D, 7

y_2D, 7