Package ‘srp’

April 24, 2019

Title Smooth-Rough Partitioning of the Regression Coefficients

Version 1.2.0

Description Performs the change-point detection in regression coefficients of linear model
by partitioning the regression coefficients into two classes of smoothness. The change-point and
the regression coefficients are jointly estimated.

Depends R (>= 3.4.0)

License GPL

LazyData true

Imports fda, mgcv

Encoding UTF-8

NeedsCompilation no

RoxygenNote 6.1.0

Author Hyeyoung Maeng [aut, cre],
  Piotr Fryzlewicz [aut]

Maintainer Hyeyoung Maeng <h.maeng@lse.ac.uk>

Repository CRAN

Date/Publication 2019-04-24 08:40:03 UTC

R topics documented:

  ncs .......................................................... 2
  predict.srp.c ........................................... 3
  predict.srp.l ........................................... 4
  sic.c ..................................................... 5
  sic.l ..................................................... 6
  srp ....................................................... 8
  srp.c ................................................... 8
  srp.l ................................................... 10
  truebeta ................................................. 11

Index 12
**ncs**

*The natural cubic spline interpolation*

**Description**

This function performs the natural cubic spline interpolation of a design matrix.

**Usage**

```r
ncs(x)
```

**Arguments**

- `x`: The design matrix.

**Details**

Usually only called by `srp.c` and `srp.l`.

**Value**

- `cf0`: The coefficient matrix for B-splines obtained by natural cubic spline interpolation of `x`.

**Author(s)**

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

**See Also**

- `srp.c`, `srp.l`

**Examples**

```r
x <- matrix(rnorm(100), ncol=10)
necs(x)$cf0
```
**predict.srp.c**

The Smooth-Rough Partition model prediction

### Description

This function performs the predictions from the results of Smooth-Rough Partition fitting.

### Usage

```r
## S3 method for class 'srp.c'
predict(object, x, ...)
```

### Arguments

- `object` An object of class either 'srp.c', returned by `srp.c`.
- `x` A new matrix you wish to fit Smooth-Rough Partition model. The dimension of row is the number of covariates.
- `...` Further parameters that can be passed to `predict.srp.c`.

### Details


### Value

- `yhat` The vector of predicted values.

### Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

### See Also

`sic.c`, `srp.c`
predict.srp.l

The Smooth-Rough Partition model prediction

Description

This function performs the predictions from the results of Smooth-Rough Partition fitting.

Usage

```r
## S3 method for class 'srp.l'
predict(object, x, ...)
```

Arguments

- `object` An object of class 'srp.l', returned by `srp.l`.
- `x` A new matrix you wish to fit Smooth-Rough Partition model. The dimension of row is the number of covariates.
- `...` Further parameters that can be passed to `predict.srp.l`.

Details


Value

- `yhat` The vector of predicted values.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

- `sic.l`
- `srp.l`
Description

This function performs the optimisation of the number of unconstrained regression parameters in Smooth-Rough Partition model by minimising SIC criterion and gives the change-point of smoothness in regression parameters.

Usage

sic.c(x.basis = x.basis, B.basis = B.basis, x = x, y = y, cf0 = cf0, maxq = maxq, fixedq = F, L = L, inisp = inisp)

Arguments

- x.basis: The b-spline basis defined for interpolated x in srp.c
- B.basis: The b-spline basis defined for constrained regression coefficient.
- x: The design matrix used in srp.c.
- y: The response variable used in srp.c.
- cf0: The coefficient matrix obtained by natural cubic spline interpolation of x in ncs.
- maxq: The maximum number of unconstrained parameters if fixedq is FALSE. Otherwise, it is considered as a unique number of unconstrained parameters.
- fixedq: If TRUE, maxq is considered as a fixed number of unconstrained parameters and if FALSE, maxq is a maximum and a sequence of possible values are investigated to select the optimal.
- L: The dimension of b-spline expansion for constrained parameters used in srp.c.
- inisp: The initial value for the optimisation of tuning parameters in srp.c.

Details

Usually only called by srp.c.

Value

The following components are obtained only when fixedq is FALSE:

- qhat: The optimal number of unconstrained parameters.
- sicq: The vector of Schwarz criterion with length maxq which is computed for the different number of unconstrained parameters.

The following components are obtained only when fixedq is TRUE:

- muhat: The estimator of constant parameter.
- bhat: The vector of evaluated constrained functional regression coefficient.
ahat The vector of unconstrained regression coefficient estimators.

etahat The vector containing both bhat and ahat with unevaluated form.

yhat The vector of estimated response variable.

sp The vector of two tuning parameters estimated by minimising generalised cross validation (GCV).

L The number of b-spline bases used for constrained regression parameters.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

sic.l, predict.srp.c, srp.c

Examples

```r
library(fda)
x <- matrix(rnorm(10000), ncol=100)
y <- matrix(rnorm(100), ncol=1)
p <- dim(x)[1] + 1
t <- seq(0, 1, length.out=dim(x)[1]*dim(x)[1])
x.basis <- as.fd(splinefun(t, x[, 1], method="natural")$basis
B.basis <- create.bspline.basis(rangeval=c(0, dim(x)[1]), norder=4, nbasis=35)
result <- sic.c(x.basis=x.basis, B.basis=B.basis, x=x, y=y, cf0=ncs(x)$cf0, maxq=10, L=35, inisp=1)
plot(result$sicq, type="b")
```

sic.l Optimisation of the (simpler) Smooth-Rough Partition model

Description

This function performs the optimisation of the number of unconstrained regression parameters in (simpler) Smooth-Rough Partition model by minimising SIC criterion and gives the change-point in regression parameters.

Usage

```r
sic.l(x.basis = x.basis, M.basis = M.basis, x = x, y = y,
       cf0 = cf0, maxq = maxq, fixedq = F)
```
Arguments

- **x.basis**: The b-spline basis defined for interpolated x in `srp.l`.
- **M.basis**: The monomial basis defined for constrained regression coefficient.
- **x**: The design matrix used in `srp.l`.
- **y**: The response variable used in `srp.l`.
- **cf0**: The coefficient matrix obtained by natural cubic spline interpolation of x in `ncs`.
- **maxq**: The maximum number of unconstrained parameters if `fixedq` is FALSE. Otherwise, it is considered as a unique number of unconstrained parameters.
- **fixedq**: If TRUE, `maxq` is considered as a fixed number of unconstrained parameters and if FALSE, `maxq` is a maximum and a sequence of possible values are investigated to select the optimal.

Details

Usually only called by `srp.l`.

Value

The following components are obtained only when `fixedq` is FALSE:

- **qhat**: The optimal number of unconstrained parameters.
- **sicq**: The vector of Schwarz criterion with length `maxq` which is computed for the different number of unconstrained parameters.

The following components are obtained only when `fixedq` is TRUE:

- **muhat**: The estimator of constant parameter.
- **bhat**: The vector of evaluated constrained functional regression coefficient.
- **ahat**: The vector of unconstrained regression coefficient estimators.
- **etahat**: The vector containing both `bhat` and `ahat` with unevaluated form.
- **yhat**: The vector of estimated response variable.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

- `sic.c`
Examples

```r
library(fda)
x <- matrix(rnorm(10000), ncol=100)
y <- matrix(rnorm(100), ncol=1)
p <- dim(x)[1] + 1
t <- seq(0, 1, length.out=dim(x)[1]*dim(x)[1])
x.basis <- as.fd(splinefun(t, x[, 1], method="natural"))$basis
M.basis <- create.monomial.basis(rangeval=c(0, dim(x)[1]), nbasis=2)
result <- sic.l(x.basis=x.basis, M.basis=M.basis, x=x, y=y, cf0=ncs(x)$cf0, maxq=10)
plot(result$sicq, type="b")
```

**srp**

*srp: Detecting the change-point of smoothness in linear regression coefficients*

**Description**

The srp package performs the change-point detection in regression coefficients of a linear model by partitioning the regression coefficients into two classes of smoothness. To start with, see the function `srp.c`.

**Author(s)**

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

**References**


**See Also**

srp.c

**srp.c**

*The Smooth-Rough Partition model fitting*

**Description**

This function performs the Smooth-Rough Partition linear regression with the input matrix.

**Usage**

```r
srp.c(x, y, maxq = max(30, ceiling(0.1 * dim(x)[1])), L = 35,
norder = 4, inisp = 1, plot = T)
```
Arguments

- **x**: A matrix you wish to fit Smooth-Rough Partition model. The dimension of row is the number of variables which are pre-ordered in terms of their importance in prediction.
- **y**: A vector you wish to use as a response variable in case of regressing y on x. If y is missing, the response variable is obtained from the last row of x.
- **maxq**: An integer specifying the maximum number of unconstrained parameters which the model can have. The default is `max(30, ceiling(0.1*dim(x)[1]))`.
- **L**: An integer specifying the dimension of b-spline expansion for the constrained (smoothed) parameters. The default is 35.
- **norder**: An integer specifying the order of b-splines. The default of 4 performs cubic splines.
- **inisp**: An initial value for optimising the tuning parameters and the default is 1.
- **plot**: If true, it gives the plot of estimated regression coefficients.

Details


Value

- **muhat**: The estimator of constant parameter.
- **bhat**: The vector of evaluated constrained functional regression coefficient.
- **ahat**: The vector of unconstrained regression coefficient estimators.
- **etahat**: The vector containing both bhat and ahat with unevaluated form.
- **yhat**: The vector of estimated response variable.
- **sic**: The vector of Schwarz criterion with length maxq which is computed for the different number of unconstrained parameters.
- **qhat**: The optimal number of unconstrained parameters selected in the model.
- **sp**: The vector of two tuning parameters estimated by minimising generalised cross validation (GCV).
- **L**: The number of bases used for constrained regression parameters.
- **norder**: The order of b-splines specified.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

- `sic.c`
- `predict.srp.c`
- `srp.l`
Examples

```r
x <- matrix(rnorm(10000), ncol=100)
srp.c(x)
```

---

**srp.l**

*The (simpler) Smooth-Rough Partition linear regression model fitting*

**Description**

This function performs same as `srp.c` except that constrained functional coefficient is estimated as a linear function.

**Usage**

```r
srp.l(x, y, maxq = max(30, ceiling(0.1 * dim(x)[1])), plot = T)
```

**Arguments**

- `x`: A matrix you wish to fit Smooth-Rough Partition model. The dimension of row is the number of variables which are pre-ordered in terms of their importance in prediction.
- `y`: A vector you wish to use as a response variable in case of regressing `y` on `x`. If `y` is missing, the response variable is obtained from the last row of `x`.
- `maxq`: An integer specifying the maximum number of unconstrained parameters which the model can have. The default is `max(30, ceiling(0.1*dim(x)[1]))`.
- `plot`: If true, it gives the plot of estimated regression coefficients.

**Details**


**Value**

- `muhat`: The estimator of constant parameter.
- `bhat`: The vector of evaluated constrained (linear) functional regression coefficient.
- `ahat`: The vector of unconstrained regression coefficient estimators.
- `etahat`: The vector containing both `bhat` and `ahat` with unevaluated form.
- `yhat`: The vector of estimated response variable.
- `sic`: The vector of Schwarz criterion with length `maxq` which is computed for the different number of unconstrained parameters.
- `qhat`: The optimal number of unconstrained parameters selected in the model.
truebeta

Author(s)
Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also
srp.c, predict.srp, sic

Examples

x <- matrix(rnorm(10000), ncol=100)
srp1(x)

truebeta

A dataset containing true regression coefficients for simulation

Description

Usage
data(truebeta)

Format
A matrix with 356 rows and 4 variables:

Case 1 true regression coefficients for case 1
Case 2 true regression coefficients for case 2
Case 3 true regression coefficients for case 3
Case 4 true regression coefficients for case 4
Index

*Topic datasets
  truebeta, 11

ncs, 2, 5, 7
predict.srp.c, 3, 3, 6, 9
predict.srp.l, 4, 4, 11
sic.c, 3, 5, 7, 9
sic.l, 4, 6, 6, 11
srp, 8
srp-package (srp), 8
srp.c, 2, 3, 5, 6, 8, 8, 10, 11
srp.l, 2, 4, 7, 9, 10

truebeta, 11