

# Package ‘edrGraphicalTools’

February 14, 2012

**Type** Package

**Title** Provide tools for dimension reduction methods

**Version** 1.0

**Date** 2010-04-14

**Author** Benoit Liquet and Jerome Saracco

**Maintainer** <benoit.liquet@isped.u-bordeaux2.fr>

**Description** This package comes to illustrate the article ‘‘A graphical tool for selecting the number of slices and the dimension of the model in SIR and SAVE approaches’’

**License** GPL (>= 2.0)

**Depends** rgl, mvtnorm

**LazyLoad** yes

**Repository** CRAN

**Date/Publication** 2010-04-27 19:09:51

## R topics documented:

edrGraphicalTools-package . . . . .	2
criterionRkh . . . . .	3
edr . . . . .	5
plot.criterionRkh . . . . .	7
plot.edr . . . . .	8
print.criterionRkh . . . . .	9
print.edr . . . . .	10
summary.criterionRkh . . . . .	10
summary.edr . . . . .	11
<b>Index</b>	<b>12</b>

---

edrGraphicalTools-package

*Provide graphical tools for dimension reduction methods*

---

## Description

This package comes to illustrate the article “A graphical tool for selecting the number of slices and the dimension of the model in SIR and SAVE approaches”. It estimate by bootstrap a square trace correlation criterion which measures the quality of the estimation of the effective dimension reduction (e.d.r) space.

## Details

Package: edrGraphicalTools  
Type: Package  
Version: 1.0  
Date: 2010-04-14  
License: GPL (>=2.0)  
LazyLoad: yes

## Author(s)

<benoit.liquet@isped.u-bordeaux2.fr>, <jerome.saracco@math.u-bordeaux1.fr>

## References

Liquet B. and Saracco, J., A graphical tool for selecting the number of slices and the dimension of the model in SIR and SAVE approaches *Revised paper*.

## See Also

[criterionRkh](#), [edr](#), [plot.criterionRkh](#)

## Examples

```
## plot example 1
set.seed(10)
n <- 500
beta <- c(1,rep(0,9))
X <- rmvnorm(n,sigma=diag(10))
eps <- rnorm(n)
Y <- (X%%beta)**3+eps*((X%%beta)**2)
## Choice a grid of integer values for H
grid.H <- c(2,5,10,15,20)
```

```
res1 <- criterionRkh(Y,X,H=grid.H,K=c(1:10),B=50,method="SIR-I")
plot(res1,choice.H=c(2,5),choice.K=c(1,2))
```

criterionRkh

*Estimation of the Rkh criterion by bootstrap method***Description**

This is the main function in the `edrGraphicalTools` package. This function estimates the square trace correlation criterion Rkh by bootstrap in order to simultaneously choose the number  $H$  of slices and the dimension  $K$  of the e.d.r space reduction. It creates objects of class `criterionRkh`. Several helper functions that require a `criterionRkh` object can then be applied to the output from this function.

**Usage**

```
criterionRkh(Y, X, H, K, indices, B = 50, method)
```

**Arguments**

<code>Y</code>	A numeric vector representing the dependent variable (a response vector).
<code>X</code>	A matrix representing the quantitative explanatory variables (bind by column).
<code>H</code>	An integer vector representing the different number of slices to be investigated. By default $H$ takes integer values in $2, 3, \dots, \text{round}(n/4)$ ; $n$ is the sample size.
<code>K</code>	An integer vector representing the different dimension $K$ to be investigated. By default $K$ takes integer values in $1, 2, \dots, \min(p, 25)$ ; $p$ is the number of explanatory variables.
<code>indices</code>	An optional integer vector specifying the indices of the bootstrap samples. By default, non-parametric resampling is performed. If used, an integer vector of length $B*n$ will be specified; ( $n$ is the length of $Y$ ).
<code>B</code>	The number of bootstrap replicates. By default $B$ equals 50.
<code>method</code>	This character string specifies the method of fitting. The options include "SIR-I", "SIR-II", and "SAVE".

**Details**

We are interested in the following semiparametric dimension reduction model proposed by Li (1991)

$$y = f(b_1'x, b_2'x, \dots, b_K'x, e)$$

where the univariate response variable  $y$  is associated with the  $p$ -dimensional regressor  $p$  only through the reduced  $K$ -dimensional variable  $(b_1'x, b_2'x, \dots, b_K'x)$  with  $K < p$ . The error term  $e$  is independent of  $x$ . The link function  $f$  and the  $b$ -vectors are unknown. We are interested in finding the linear subspace spanned by the  $K$  unknown  $b$ -vector, called the effective dimension reduction (e.d.r.) space. We focus on the SIR (named SIR-I), SIR-II and SAVE methods to estimate the e.d.r space. The slicing step of these methods depends on the number  $H$  of slices. We propose a

naive bootstrap estimation of the square trace correlation criterion to allow selection of an “optimal” number of slices and to simultaneously select the corresponding suitable dimension  $K$  (number of the linear combinations of  $x$ ).

### Value

criterionRkh returns an object that inherits from criterionRkh (the name of the type is the value of the method argument), with attributes:

Rkhbootmean	A matrix corresponding of the estimation by bootstrap of the square trace criterion Rkh (h in rows and k in columns).
Rkhboot	A list including the result of the estimation of the square trace criterion for each bootstrap replicate.
method	the dimension reduction method used.
n	Number of subject.
H	A vector representing the different numbers H of slices investigated.
K	A vector representing the different dimensions K investigated.
indices	An vector of integers representing the indices of the bootstrap sample used.

### Author(s)

Benoit Liquet, <benoit.liquet@isped.u-bordeaux2.fr> and Jerome Saracco <jerome.saracco@math.u-bordeaux1.fr>

### References

Liquet B. and Saracco, J., A graphical tool for selecting the number of slices and the dimension of the model in SIR and SAVE approaches. *Revised paper*.

Li, K.C. (1991). Sliced inverse regression for dimension reduction, with discussions. *Journal of the American Statistical Association* 86, 316-342.

Cook, R. D. and Weisberg, S. (1991). Discussion of “Sliced inverse regression”. *Journal of the American Statistical Association*, 86, 328-332.

### See Also

[edr](#), [summary.criterionRkh](#), [plot.criterionRkh](#)

### Examples

```
## simulated example 1
set.seed(10)
n <- 500
beta <- c(1,rep(0,9))
X <- rmvnorm(n,sigma=diag(10))
eps <- rnorm(n)
Y <- (X%*%beta)**3+eps*((X%*%beta)**2)
## Choice a grid of values for H
grid.H <- c(2,5,10,15,20,30)
```

```

res1 <- criterionRkh(Y,X,H=grid.H,B=50,method="SIR-I")
res1
plot(res1,choice.H=c(2,5),choice.K=c(1,2))
## Estimation for SIR-II method with the same bootstrap replicate than for SIR-I
res2 <- criterionRkh(Y,X,H=grid.H,indices=res1$indices,B=50,method="SIR-II")
res2

```

edr

*Main function for estimation of the e.d.r space***Description**

It creates objects of class `edr` to estimate the effective dimension regression (e.d.r) space. Several helper functions that require a `edr` object can then be applied to the output from this function.

**Usage**

```
edr(Y, X, H, K, method)
```

**Arguments**

<code>Y</code>	A numeric vector representing the dependent variable (a response vector).
<code>X</code>	A matrix representing the quantitative explanatory variables (bind by column).
<code>H</code>	The chosen number of slices.
<code>K</code>	The chosen dimension $K$ .
<code>method</code>	This character string specifies the method of fitting. The option includes "SIR-I", "SIR-II", and "SAVE".

**Details**

We interested in the following semiparametric dimension reduction model proposed by Li (1991)

$$y = f(b_1'x, b_2'x, \dots, b_K'x, e)$$

where the univariate response variable  $y$  is associated with the  $p$ -dimensional regressor  $p$  only through the reduced  $K$ -dimensional variable  $(b_1'x, b_2'x, \dots, b_K'x)$  with  $K < p$ . The error term  $e$  is independent of  $x$ . The link function  $f$  and the  $b$ -vectors are unknown. We are interested in finding the linear subspace spanned by the  $K$  unknown  $b$ -vector, called the effective dimension reduction (e.d.r.) space. We focus on the SIR, SIR-II and SAVE methods to estimate the e.d.r space. The slicing step of these methods depends on the number  $H$  of slices. We propose with the function `criterionRkh` a naive bootstrap estimation of the square trace correlation criterion to allow selection of an "optimal" number  $H$  of slices and simultaneously the corresponding suitable dimension  $K$  (number of the linear combination of  $x$ ). After choosing an optimal couple  $(H, K)$  for the best estimation method (the square trace correlation criterion closest to one), the e.d.r space could be estimate with this function. Each method consists in a spectral decomposition of a matrix of interest. The eigenvectors of this matrix associated of the  $K$  largest eigenvalues are e.d.r directions.

**Value**

edr returns an object that inherits from `edr`, with attributes:

<code>matEDR</code>	A matrix corresponding of the eigenvectors of the interest matrix
<code>eigvalEDR</code>	The eigenvalues of the matrix of interest
<code>K</code>	The chosen dimension.
<code>H</code>	The chosen number of slices.
<code>n</code>	Sample size.
<code>method</code>	The dimension reduction method used.
<code>X</code>	The matrix of the quantitative explanatories variables (bind by column).
<code>Y</code>	The numeric vector of the dependent variable (a response vector).

**Author(s)**

Benoit Liquet, <benoit.liquet@isped.u-bordeaux2.fr> and Jerome Saracco <jerome.saracco@math.u-bordeaux1.fr>

**References**

Liquet B. and Saracco, J., A graphical tool for selecting the number of slices and the dimension of the model in SIR and SAVE approaches. *Revised paper*.

Li, K.C. (1991). Sliced inverse regression for dimension reduction, with discussions. *Journal of the American Statistical Association* 86, 316-342.

Cook, R. D. and Weisberg, S. (1991). Discussion of "Sliced inverse regression". *Journal of the American Statistical Association*, 86, 328-332.

**See Also**

[criterionRkh](#), [summary.edr](#), [plot.edr](#)

**Examples**

```
set.seed(10)
n <- 500
beta1 <- c(1,1,rep(0,8))
beta2 <- c(0,0,1,1,rep(0,6))
X <- rmvnorm(n,sigma=diag(1,10))
eps <- rnorm(n)
Y <- (X%*%beta1)**2+(X%*%beta2)**2+eps

## Estimation of the trace square criterion
## grid.H <- c(2,5,10,15,20,30)
## res2 <- criterionRkh(Y,X,H=grid.H,B=50,method="SIR-II")
## summary(res2)
## plot(res2)

## Estimation of the e.d.r direction for K=2 and H=2 and SIR-II method
edr2 <- edr(Y,X,H=2,K=2,method="SIR-II")
```

```
summary(edr2)
plot(edr2)
```

---

plot.criterionRkh      *Graphical tools for the bootstrap criterion*

---

## Description

We propose a 3D-graphical tool which can be useful to select the suitable couple (H,K). We also add some boxplots of the distribution of the bootstrap criterion.

## Usage

```
## S3 method for class 'criterionRkh'
plot(x, choice.H, choice.K, ...)
```

## Arguments

x	The name of an object of class <code>criterionRkh</code> .
choice.H	An optional integer vector specifying the values of H for which a boxplot of the distribution of the bootstrap criterion is plotted for all the values of K specified in the <code>criterionRkh</code> object. The values of <code>choice.H</code> must be included in the values of H specified in the <code>criterionRkh</code> object.
choice.K	An optional integer vector specifying the values of K for which a boxplot of the distribution of the bootstrap criterion is plotted for all the values of H specified in the <code>criterionRkh</code> object. The values of <code>choice.K</code> must be included in the values of K specified in the <code>criterionRkh</code> object.
...	Some methods for this generic require additional arguments. None are used in this method.

## Value

Returns graphs.

## Author(s)

Benoit Liquet, <benoit.liquet@isped.u-bordeaux2.fr> and Jerome Saracco <jerome.saracco@math.u-bordeaux1.fr>

## References

Liquet B. and Saracco, J., A graphical tool for selecting the number of slices and the dimension of the model in SIR and SAVE approaches. *Revised paper*

## See Also

[criterionRkh](#)

**Examples**

```
## see example in function criterionRkh
```

---

```
plot.edr
```

```
Basic plot of a edr object
```

---

**Description**

For an edr object with attributes  $K \leq 2$ , this function plots the response  $y$  versus each new estimate indice with an estimation of the link function. For  $K = 2$ , a 3D plot of  $y$  versus the two estimate indices is represented. A smooth estimate of the link function is also represented. For  $K > 2$ , the pairs function is used.

**Usage**

```
## S3 method for class 'edr'
plot(x, ...)
```

**Arguments**

<code>x</code>	The name of an object of class <code>edr</code> .
<code>...</code>	Some methods for this generic require additional arguments. None are used in this method.

**Value**

Returns graphs.

**Author(s)**

Benoit Liquet, <benoit.liquet@isped.u-bordeaux2.fr> and Jerome Saracco <jerome.saracco@math.u-bordeaux1.fr>

**See Also**

[summary.edr](#)

**Examples**

```
## simulated example
set.seed(10)
n <- 500
beta1 <- c(1,1,rep(0,8))
beta2 <- c(0,0,1,1,rep(0,6))
X <- rmvnorm(n,sigma=diag(1,10))
eps <- rnorm(n)
Y <- (X%*%beta1)**2+(X%*%beta2)**2+eps
edr2 <- edr(Y,X,H=2,K=2,method="SIR-II")
```

```
plot(edr2)
## edr4 <- edr(Y,X,H=2,K=4,method="SIR-II")
## plot(edr4)
```

---

`print.criterionRkh`     *Print a Summary of a criterionRkh Object*

---

### **Description**

This is a method for the function `print` for objects of the class `criterionRkh`.

### **Usage**

```
## S3 method for class 'criterionRkh'
print(x, ...)
```

### **Arguments**

`x`                    An object of class `criterionRkh` generated by the function [criterionRkh](#).  
`...`                Further arguments passed to or from other methods.

### **Value**

A summary of the `criterionRkh` Object is returned.

### **Author(s)**

Benoit Liquet, <benoit.liquet@isped.u-bordeaux2.fr> and Jerome Saracco <jerome.saracco@math.u-bordeaux1.fr>

### **See Also**

[criterionRkh](#)

print.edr                      *Print a Summary of a edr Object*

---

### Description

This is a method for the function print for objects of the class edr.

### Usage

```
## S3 method for class 'edr'  
print(x, ...)
```

### Arguments

x	An object of class edr generated by the function <a href="#">edr</a> .
...	Further arguments to be passed to or from other methods. They are ignored in this function

### Value

A summary of the edr Object is returned.

### Author(s)

Benoit Liquet, <benoit.liquet@isped.u-bordeaux2.fr> and Jerome Saracco <jerome.saracco@math.u-bordeaux1.fr>

### See Also

[edr](#)

---

summary.criterionRkh    *Print a Summary of a criterionRkh Object*

---

### Description

This is a method for the function summary for objects of the class criterionRkh.

### Usage

```
## S3 method for class 'criterionRkh'  
summary(object, ...)
```

**Arguments**

object            An object of class `critereionRkh` generated by the function `critereionRkh`.  
...                Further arguments passed to or from other methods.

**Value**

A summary of the `critereionRkh` Object is returned.

**Author(s)**

Benoit Liquet, <benoit.liquet@isped.u-bordeaux2.fr> and Jerome Saracco <jerome.saracco@math.u-bordeaux1.fr>

**See Also**

[critereionRkh](#)

---

summary.edr

*Print a Summary of a edr Object*

---

**Description**

This is a method for the function `summary` for objects of the class `edr`.

**Usage**

```
## S3 method for class 'edr'  
summary(object, ...)
```

**Arguments**

object            An object of class `edr` generated by the function `edr`.  
...                further arguments passed to or from other methods.

**Value**

A summary of the `edr` Object is returned.

**Author(s)**

Benoit Liquet, <benoit.liquet@isped.u-bordeaux2.fr> and Jerome Saracco <jerome.saracco@math.u-bordeaux1.fr>

**See Also**

[edr](#)

# Index

- \*Topic **package**
  - edrGraphicalTools-package, 2
- \*Topic **print**
  - print.criterionRkh, 9
  - print.edr, 10
- \*Topic **regression**
  - criterionRkh, 3
  - edr, 5
  - plot.criterionRkh, 7
  - plot.edr, 8
  - summary.criterionRkh, 10
  - summary.edr, 11

criterionRkh, 2, 3, 5–7, 9, 11

edr, 2, 4, 5, 8, 10, 11

edrGraphicalTools  
(edrGraphicalTools-package), 2

edrGraphicalTools-package, 2

plot.criterionRkh, 2, 4, 7

plot.edr, 6, 8

print.criterionRkh, 9

print.edr, 10

summary.criterionRkh, 4, 10

summary.edr, 6, 8, 11