Package ‘nsdr’

June 3, 2021

Type Package

Title Nonlinear Sufficient Dimension Reduction

Version 0.1.1

Description Provides tools to implement both unsupervised and supervised nonlinear dimension reduction methods. Principal Component Analysis (PCA), Sliced Inverse Regression (SIR), and Sliced Average Variance Estimation (SAVE) are useful methods to reduce the dimensionality of covariates. However, they produce linear combinations of covariates. Kernel PCA, generalized SIR, and generalized SAVE address this problem by extending the applicability of the dimension reduction problem to nonlinear settings. This package includes a comprehensive algorithm for kernel PCA, generalized SIR, and generalized SAVE, including methods for choosing tuning parameters and some essential functions.

Depends R (>= 3.5.0)

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1.9001

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Maintainer Kyongwon Kim <kimk@ewha.ac.kr>

Author Bing Li [aut],
Kyongwon Kim [aut, cre]

Repository CRAN

Date/Publication 2021-06-03 06:10:02 UTC

R topics documented:

`gcv` *gcv*

Description

gcv

Usage

```
gcv(x,y,eps,which,ytype,complex.x,complex.y)
```

Arguments

`x`	input predictor matrix from training set
`y`	response variables
`eps`	candidate
`which`	choose between ex and ey
`ytype`	type of response variables
`complex.x`	tuning parameter for the Gaussian kernel in X
`complex.y`	tuning parameter for the Gaussian kernel in Y

Value

gcv criterion

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
n = 50; p = 5; sigma = 1;
x = matrix(rnorm(n*p),n,p) ; err = rnorm(n)
y = (x[,1]+1)^2 + sigma*err; ex=0.01 ; ey=0.01; candidate=0.01
epsx <- gcv(x,y,candidate,"ex", "categorical",1,1)
epsy <- gcv(x,y,candidate,"ey", "categorical",1,1)
```

---

gram.dis *gram.dis*

Description

gram.dis

Usage

```
gram.dis(y)
```

Arguments

y	discrete vector

Value

gram matrix for discrete vector

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
toy <- c(1,2,3,1)
result=gram.dis(toy)
```

gram.gauss *gram.gauss*

Description

gram.gauss

Usage

```
gram.gauss(x,x.new,complexity)
```

Arguments

`x`	previous observations
`x.new`	new observations
`complexity`	tuning parameter in Gaussian kernel

Value

Gram matrix for the Gaussian kernel. This also can be used to project predictor on the testing set.

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
old <- matrix(c(1,2,3,4),2,2)
new <- matrix(c(5,6,7,8),2,2)
result <- gram.gauss(old,new,1)
```

gramx *gramx*

Description

gramx

Usage

```
gramx(x,complexity)
```

Arguments

`x`	data
`complexity`	tuning parameter in Gaussian kernel

Value

gram matrix Q x KX x Q

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
vec <- matrix(rnorm(4),2,2)
res <- gramx(vec,1)
```

---

gsave *gsave*

Description

gsave

Usage

```
gsave(x,x.new,y,ytype,ex,ey,comx,comy,r)
```

Arguments

`x`	input predictor matrix from training set
`x.new`	input predictor matrix from testing set
`y`	response variables
`ytype`	type of response variables
`ex`	tuning parameter for the Tychonoff reguralized inverse for GX
`ey`	tuning parameter for the Tychonoff reguralized inverse for GY
`comx`	tuning parameter for the Gaussian kernel in X
`comy`	tuning parameter for the Gaussian kernel in Y
`r`	number of dimension

Value

pred: sufficient predictors from GSAVE

obj.mat: objective matrix of GSAVE

eig.val: the first r eigenvalues from the eigendecomposition of the objective matrix

eig.vec: the first r eigenvectors from the eigendecomposition of the objective matrix

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
n = 50; p = 5; sigma = 1;
x = matrix(rnorm(n*p),n,p) ; err = rnorm(n)
y = x[,1]/(0.5+(x[,1]+1)^2) + sigma*err; ex=0.01 ; ey=0.01
gsave_res <- gsave(x,x,y,"scalar",ex,ey,1,1,1)
```

---

gsir *gsir*

Description

gsir

Usage

```
gsir(x,x.new,y,ytype,ex,ey,complex.x,complex.y,r)
```

Arguments

`x`	input predictor matrix from training set
`x.new`	input predictor matrix from testing set
`y`	response variables
`ytype`	type of response variables
`ex`	tuning parameter for the Tychonoff regularized inverse for GX
`ey`	tuning parameter for the Tychonoff regularized inverse for GY
`complex.x`	tuning parameter for the Gaussian kernel in X
`complex.y`	tuning parameter for the Gaussian kernel in Y
`r`	number of dimension

Value

suff.pred: sufficient predictors from GSIR

obj.mat: objective matrix of GSIR

eig.val: the first r eigenvalues from the eigendecomposition of the objective matrix

eig.vec: the first r eigenvectors from the eigendecomposition of the objective matrix

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
n = 50; p = 5; sigma = 1;
x = matrix(rnorm(n*p),n,p) ; err = rnorm(n)
y = sin(0.5+(x[,1]+1)^2) + sigma*err; ex=0.01 ; ey=0.01
gsir_res <- gsir(x,x,y,"scalar",ex,ey,1,1,1)
```

`kpca` *kpca*

Description

kpca

Usage

```
kpca(x,complexity)
```

Arguments

`x`	dataset
`complexity`	tuning parameter in Gaussian kernel. larger complexity means a wiggly kernel function

Value

principal component

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
n = 50; p = 5
x = matrix(rnorm(n*p),n,p)
pred=kpca(x,1)[,1:3]
```

`matpower` *matpower*

Description

matpower

Usage

```
matpower(a,alpha)
```

Arguments

`a`	matrix
`alpha`	power

Value

power of a matrix

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
mat <- matrix(rnorm(4,1,3),2,2)
invmat <- matpower(mat,-1)
```

`mppower` *mppower*

Description

mppower

Usage

```
mppower(matrix,power,ignore)
```

Arguments

`matrix`	input matrix
`power`	power
`ignore`	ignoring criterion

Value

Moore penrose inverse

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
a <- matrix(rnorm(4,0,1),2,2)
mppower(a,-1,0.2)
```

`onorm` *onorm*

Description

onorm

Usage

```
onorm(a)
```

Arguments

`a`	input matrix

Value

result of the operator norm

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
a <- matrix(c(1,2,3,4),2,2)
result <- onorm(a)
```

`pendigits.tes` *Pen-Based Recognition of Handwritten Digits Data Set (testing dataset)*

Description

The data is about the recognition of handwritten numbers from 0 to 9. There are 30 writers in the training dataset and each participant are asked to write 250 digits in random order. Without missing data, this dataset has 7494 observations. The experiment uses WACOM tablet, which has 500 x 500 pixel resolutions and normalized it to a maximum scale of 100. The researcher considers spatial resampling. Thus, for each digit, eight pairs of 2 dimensional (x axis and y axis) locations are recorded, which makes this dataset have 16 dimensional predictor variables.

Usage

```
pendigits.tes
```

Format

A data frame with 2219 observations on 17 variables.

The first column to the 16th column represent resampled values of the pairs of points. The 17th column presents the digit (0 to 9).

Source

https://archive.ics.uci.edu/ml/datasets/Pen-Based+Recognition+of+Handwritten+Digits

References

F. Alimoglu (1996) Combining Multiple Classifiers for Pen-Based Handwritten Digit Recognition, MSc Thesis, Institute of Graduate Studies in Science and Engineering, Bogazici University. F. Alimoglu, E. Alpaydin, "Methods of Combining Multiple Classifiers Based on Different Representations for Pen-based Handwriting Recognition," Proceedings of the Fifth Turkish Artificial Intelligence and Artificial Neural Networks Symposium (TAINN 96), June 1996, Istanbul, Turkey.

`pendigits.tra` *Pen-Based Recognition of Handwritten Digits Data Set (training dataset)*

Description

The data is about the recognition of handwritten numbers from 0 to 9. There are 30 writers in the training dataset and each participant are asked to write 250 digits in random order. Without missing data, this dataset has 7494 observations. The experiment uses WACOM tablet, which has 500 x 500 pixel resolutions and normalized it to a maximum scale of 100. The researcher considers spatial resampling. Thus, for each digit, eight pairs of 2 dimensional (x axis and y axis) locations are recorded, which makes this dataset have 16 dimensional predictor variables.

Usage

```
pendigits.tra
```

Format

A data frame with 2219 obsevations on 17 variables.

The first column to the 16th column represent resampled values of the pairs of points. The 17th column presents the digit (0 to 9).

Source

https://archive.ics.uci.edu/ml/datasets/Pen-Based+Recognition+of+Handwritten+Digits

References

F. Alimoglu (1996) Combining Multiple Classifiers for Pen-Based Handwritten Digit Recognition, MSc Thesis, Institute of Graduate Studies in Science and Engineering, Bogazici University. F. Alimoglu, E. Alpaydin, "Methods of Combining Multiple Classifiers Based on Different Representations for Pen-based Handwriting Recognition," Proceedings of the Fifth Turkish Artificial Intelligence and Artificial Neural Networks Symposium (TAINN 96), June 1996, Istanbul, Turkey.

`ridgepower` *ridgepower*

Description

ridgepower

Usage

```
ridgepower(a,e,c)
```

Arguments

`a`	square matrix
`e`	tuning parameter
`c`	power

Value

matrix with the power

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
x <- matrix(c(1:4),2,2)
result <- ridgepower(x, 0.001, -1)
```

`spearman` *standmat*

Description

standmat

Usage

```
spearman(x1,x2)
```

Arguments

x1	first argument
x2	second argument

Value

standardized matrix

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
x1 <- rnorm(100)
x2 <- rnorm(100)
spearman(x1,x2)
```

`standmat` *standmat*

Description

standmat

Usage

```
standmat(x)
```

Arguments

x	matrix

Value

standardized matrix

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
mat <- matrix(rnorm(4),2,2)
standmat(mat)
```

`sym` *sym*

Description

sym

Usage

```
sym(a)
```

Arguments

`a`	any matrix

Value

symmetrize matrix when matrix is theoretically symmetric but not in numerically

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
ex <- matrix(c(1.1,2.1,1.2,2.2),2,2)
result <- sym(ex)
```

`tr` *trace*

Description

trace

Usage

```
tr(a)
```

Arguments

`a`	any matrix

Value

trace value of the matrix

References

Li, B. (2018). Sufficient dimension reduction: Methods and applications with R. CRC Press.

Examples

```
mat <- matrix(2,2,2)
tr(mat)
```

`wine` *Chemical ingredients of wine dataset*

Description

This dataset is about chemical analysis of 178 wines in a particular region of Italy. There are three cultivars with 59, 71, and 48, respectively.

Usage

```
wine
```

Format

A data frame with 178 observations on 14 variables.

Attributes 1. cultivar 2. alcohol 3. malic acid 4. ash 5. alcalinity of ash 6. magnesium 7. total phenols 8. flavanoids 9. nonflavanoid phenols 10. oroanthocyanins 11. color intensity 12. hue 13. OD280/OD315 of diluted wines 14. proline.

Source

https://archive.ics.uci.edu/ml/datasets/wine

References

Forina, M., Leardi, R., Armanino, C., Lanteri, S., Conti, P., & Princi, P. (1988). PARVUS: An extendable package of programs for data exploration, classification and correlation. Journal of Chemometrics, 4(2), 191-193..

Index