

Package ‘OPC’

January 20, 2025

Title The Online Principal Component Estimation Method

Date 2023-05-15

Version 0.0.2

Description The online principal component method can process the online data set. The philosophy of the package is described in Guo G. (2018) <[doi:10.1080/10485252.2018.1531130](https://doi.org/10.1080/10485252.2018.1531130)>.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.2.3

Imports MASS, stats

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Chunjie Wei [aut, cre],
Guangbao Guo [aut]

Maintainer Chunjie Wei <wcj1236@163.com>

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2023-05-15 14:10:02 UTC

Contents

Cloud	2
HTRU	3
IPC	4
PPC	4
SAPC	5
Wine	6
Index	8

Cloud

Cloud

Description

A data frame with 1024 observations on the following 10 variables.

Usage

```
data("Cloud")
```

Format

A data frame with 1024 observations on the following 10 variables.

x1 a numeric vector
x2 a numeric vector
x3 a numeric vector
x4 a numeric vector
x5 a numeric vector
x6 a numeric vector
x7 a numeric vector
x8 a numeric vector
x9 a numeric vector
x10 a numeric vector

Details

The data sets we propose to analyse are constituted of 1024 vectors, each vector includes 10 parameters. You can think of it as a 1024*10 matrix.

Source

The Cloud data set comes from the UCI database.

References

NA

Examples

```
data(Cloud)  
## maybe str(Cloud) ; plot(Cloud) ...
```

HTRU

HTRU

Description

A data frame with 10000 observations on the following 9 variables.

Usage

```
data("HTRU")
```

Format

A data frame with 10000 observations on the following 9 variables.

A a numeric vector

B a numeric vector

C a numeric vector

D a numeric vector

E a numeric vector

F a numeric vector

G a numeric vector

H a numeric vector

I a numeric vector

Details

Pulsar candidates collected during the HTRU survey. Pulsars are a type of star, of considerable scientific interest. Candidates must be classified in to pulsar and non-pulsar classes to aid discovery.

Source

The HTRU data set comes from the UCI database.

References

NA

Examples

```
data(HTRU)
## maybe str(HTRU) ; plot(HTRU) ...
```

IPC *The incremental principal component method can handle online data sets.*

Description

The incremental principal component method can handle online data sets.

Usage

```
IPC(data, m, eta)
```

Arguments

data is an online data set
 m is the number of principal component
 eta is the proportion of online data to total data

Value

T2,T2k,V,Vhat,lambdahat,time

Examples

```
library(MASS)
n=2000;p=20;m=9;
mu=t(matrix(rep(runif(p,0,1000),n),p,n))
mu0=as.matrix(runif(m,0))
sigma0=diag(runif(m,1))
F=matrix(mvrnorm(n,mu0,sigma0),nrow=n)
A=matrix(runif(p*m,-1,1),nrow=p)
D=as.matrix(diag(rep(runif(p,0,1))))
epsilon=matrix(mvrnorm(n,rep(0,p),D),nrow=n)
data=mu+F%*%t(A)+epsilon
IPC(data=data,m=m,eta=0.8)
```

PPC *The perturbation principal component method can handle online data sets.*

Description

The perturbation principal component method can handle online data sets.

Usage

```
PPC(data, m, eta)
```

Arguments

data is an online data set
 m is the number of principal component
 eta is the proportion of online data to total data

Value

T2,T2k,V,Vhat,lambdahat,time

Examples

```
library(MASS)
n=2000;p=20;m=9;
mu=t(matrix(rep(runif(p,0,1000),n),p,n))
mu0=as.matrix(runif(m,0))
sigma0=diag(runif(m,1))
F=matrix(mvrnorm(n,mu0,sigma0),nrow=n)
A=matrix(runif(p*m,-1,1),nrow=p)
D=as.matrix(diag(rep(runif(p,0,1))))
epsilon=matrix(mvrnorm(n,rep(0,p),D),nrow=n)
data=mu+F%%t(A)+epsilon
PPC(data=data,m=m,eta=0.8)
```

SAPC

The stochastic approximate component method can handle online data sets.

Description

The stochastic approximate component method can handle online data sets.

Usage

SAPC(data, m, eta, alpha)

Arguments

data is a online data set
 m is the number of principal component
 eta is the proportion of online data to total data
 alpha is the step size

Value

T2,T2k,V,Vhat,lambdahat,time

Examples

```

library(MASS)
n=2000;p=20;m=9;
mu=t(matrix(rep(runif(p,0,1000),n),p,n))
mu0=as.matrix(runif(m,0))
sigma0=diag(runif(m,1))
F=matrix(mvrnorm(n,mu0,sigma0),nrow=n)
A=matrix(runif(p*m,-1,1),nrow=p)
D=as.matrix(diag(rep(runif(p,0,1))))
epsilon=matrix(mvrnorm(n,rep(0,p),D),nrow=n)
data=mu+F%*%t(A)+epsilon
SAPC(data=data,m=m,eta=0.8,alpha=1)

```

Wine

Wine

Description

A data frame with 177 observations on the following 13 variables.

Usage

```
data("Wine")
```

Format

A data frame with 177 observations on the following 13 variables.

X14.23 a numeric vector

X1.71 a numeric vector

X2.43 a numeric vector

X15.6 a numeric vector

X127 a numeric vector

X2.8 a numeric vector

X3.06 a numeric vector

X.28 a numeric vector

X2.29 a numeric vector

X5.64 a numeric vector

X1.04 a numeric vector

X3.92 a numeric vector

X1065 a numeric vector

Details

These data are the results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The analysis determined the quantities of 13 constituents found in each of the three types of wines.

Source

The Wine data set comes from the UCI database.

References

NA

Examples

```
data(Wine)
## maybe str(Wine) ; plot(Wine) ...
```

Index

* datasets

Cloud, [2](#)

HTRU, [3](#)

Wine, [6](#)

Cloud, [2](#)

HTRU, [3](#)

IPC, [4](#)

PPC, [4](#)

SAPC, [5](#)

Wine, [6](#)