Package 'xyz'

April 3, 2017

Type Package

Title The 'xyz' Algorithm for Fast Interaction Search in
High-Dimensional Data

Version 0.2

Date 2016-10-17

Maintainer Gian-Andrea Thanei <thanei@stat.math.ethz.ch>

Author Gian-Andrea Thanei

Description High dimensional interaction search by brute force requires a
quadratic computational cost in the number of variables. The xyz algorithm prov-
ably finds strong interactions in almost linear time.
For details of the algorithm see: G. Thanei, N. Meinshausen and R. Shah (2016). The xyz algo-
rithm for fast interaction search in high-
dimensional data <https://arxiv.org/pdf/1610.05108v1.pdf>.

License GPL

LazyData TRUE

Imports stats, Rcpp (>= 0.12.6)

LinkingTo Rcpp

RoxygenNote 6.0.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

SystemRequirements C++11

NeedsCompilation yes

Repository CRAN

Date/Publication 2017-04-03 16:39:20 UTC

R topics documented:

---

xyz	*xyz: A package for fast interaction search in high dimensional data using the xyz algorithm.*

Description

xyz: A package for fast interaction search in high dimensional data using the xyz algorithm.

xyz functions

xyz_search, xyz_regression

References

G. Thanei, N. Meinshausen and R. Shah (2016). The xyz algorithm for fast interaction search in high-dimensional data. <https://arxiv.org/pdf/1610.05108v1.pdf>

---

xyz_regression	*Elasticnet with interactions (glmnet)*

Description

Elasticnet with interactions (glmnet)

Usage

```
xyz_regression(X, Y, lambdas = NULL, n_lambda = 10, alpha = 0.9, L = 10,
  standardize = TRUE, standardize_response = TRUE)
```

Arguments

X	A matrix.
Y	A vector.
lambdas	A vector of decreasing real numbers containing user specified values of lambda.
n_lambda	A natural number indicating how long the path of lambdas should be.
alpha	A real number between 0 and 1 (the elastic net parameter)
L	An integer indicating how many projection steps are performed.
standardize	A boolean indicating if X should be scaled and centered.
standardize_response	
	A boolean indicating if Y should be scaled and centered.

Value

N strongest interactions (of type `type`) between X and Y after L projections.

References

G. Thanei, N. Meinshausen and R. Shah (2016). The xyz algorithm for fast interaction search in high-dimensional data. <https://arxiv.org/pdf/1610.05108v1.pdf>

Examples

```
n<-300
p<-1000
#build matrix of predictors
X<-matrix(rnorm(n*p),n,p)
#build a main effect and an interaction into Y
Y<-4*X[,1]*X[,2]-5*X[,4]+rnorm(n)
result<-xyz_regression(X,Y,n_lambda=10,alpha=0.9,L=10)
#print the result
print(result)
#plot the result
plot(result)
```

---

xyz_search   *Interaction search*

Description

Interaction search

Usage

```
xyz_search(X, Y, L = 10, N = 100, binary = TRUE, negative = TRUE)
```

Arguments

X	A matrix.
Y	A vector.
L	An integer indicating how many projection steps are performed.
N	A integer, controlling the number of pairs that will be returned in the end.
binary	A logical indicating if X is binary or continuous.
negative	A logical indicating if also negative interactions should be searched for.

Value

N strongest interactions between X and Y after L projections.

References

G. Thanei, N. Meinshausen and R. Shah (2016). The xyz algorithm for fast interaction search in high-dimensional data. <https://arxiv.org/pdf/1610.05108v1.pdf>

Examples

```
n<-300
p<-1000
#construct a binary matrix
X<-matrix(sample(c(-1,1),replace=TRUE,n*p),n,p)
#set an interaction of the pair (1,2)
Y<-X[,1]*X[,2]+rnorm(n)
#run the interaction search
result<-xyz_search(X,Y,L=10,N=10,binary=TRUE,negative=TRUE)
#print the result
print(result)
```

Index