Package 'SEPaLS'

January 20, 2025

Title Shrinkage for Extreme Partial Least-Squares (SEPaLS)

Date 2023-10-11

Type Package

Version 0.1.0

Description Regression context for the Partial Least Squares framework for Extreme values. Estimations of the Shrinkage for Extreme Partial Least-Squares (SEPaLS) estimators, an adaptation of the original Partial Least Squares (PLS) method tailored to the extreme-value framework.
The SEPaLS project is a joint work by Stephane Girard, Hadrien Lorenzo and Julyan Arbel.
R code to replicate the results of the paper is available at <https://github.com/hlorenzo/SEPaLS_simus>.
Extremes within PLS was already studied by one of the authors, see M Bousebeta, G Enjolras, S Girard (2023) <doi:10.1016/j.jmva.2022.105101>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.2.0

NeedsCompilation no

Author Stephane Girard [aut],
Julyan Arbel [aut],
Hadrien Lorenzo [aut, cre, cph]

Maintainer Hadrien Lorenzo <hadrien.lorenzo@univ-amu.fr>

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2023-10-24 18:20:06 UTC

Contents

`bootstrap.SEPaLS` *Bootstrap function for SEPaLS estimator.*

Description

Bootstrap function for SEPaLS estimator.

Usage

```
bootstrap.SEPaLS(
  X,
  Y,
  yn,
  type = c("vMF", "Laplace"),
  mu0 = NULL,
  kappa0 = NULL,
  lambda = NULL,
  B = 20
)
```

Arguments

`X`	$(n \times p)$-dimensional matrix of the covariates.
`Y`	(n)-dimensional vector of the response.
`yn`	y_n the quantile corresponding to lowest values of Ys to put in the tail.
`type`	character, whether `vMF` for von Mises-Fisher prior or `Laplace` for Laplace prior. See details.
`mu0`	μ_0, unitary (p)-dimensional vector. The direction parameter for the `vMF` prior.
`kappa0`	κ_0, positive. The concentration parameter for the `vMF` prior.
`lambda`	λ, positive. The concentration parameter for the `Laplace` prior.
`B`	positive integer. The number of bootstrap samples on which estimate the SEPaLS directions. Default to 20.

Value

A list with two elements:

- `ws`: A $(B \times p)$-dimensional matrix with each row corresponding to the *SEPaLS* direction estimated on each bootstrap sample.
- `cor`: The correlation of each estimate direction on the Out-Of-Bag (OOB) sample with the response.

See Also

`SEPaLS`

Examples

```
set.seed(5)
n <- 3000
p <- 10
X <- matrix(rnorm(n*p),n,p)
beta <- c(5:1,rep(0,p-5)) ; beta <- beta/sqrt(sum(beta^2))
Y <- (X%*%beta)^3 + rnorm(n)
boot.sepals_Laplace <- bootstrap.SEPaLS(X,Y,yn=1,type="Laplace",lambda=0.01,
B=100)
boxplot(boot.sepals_Laplace$ws);abline(h=0,col="red",lty=2)
```

`maximum_Likelihood_SEPaLS`

Maximum Likelihood estimator

Description

Maximum Likelihood estimator

Usage

```
maximum_Likelihood_SEPaLS(X, Y, yn)
```

Arguments

`X`	$(n \times p)$-dimensional matrix of the covariates.
`Y`	(n)-dimensional vector of the response.
`yn`	the quantile corresponding to the lowest values of Ys to put in the tail.

Value

The maximum likelihood estimator.

Examples

```
n <- 3000
p <- 10
X <- matrix(rnorm(n*p),n,p)
beta <- c(5:1,rep(0,p-5)) ; beta <- beta/sqrt(sum(beta^2))
Y <- X%*%beta + rnorm(n,sd=1/3)
estimators <- do.call(rbind,lapply(seq(0,1,length.out=100),function(pp){
  yn <- quantile(Y,probs = pp)
  maximum_Likelihood_SEPaLS(X,Y,yn)
}))
matplot(estimators,type="l",lty=1,col=c(rep(2,5),rep(1,p-5)))
abline(h=beta/sqrt(sum(beta^2)),col=c(rep(2,5),rep(1,p-5)))
```

ricaCarrots — *The RICA dataset describing the production of carrots (open field) (in quintals) from 2000 to 2015.*

Description

A subset of data from the 'agreste' French governmental website <https://agreste.agriculture.gouv.fr/agreste-web/servicon/I.2/listeTypeServicon/>.

Usage

```
data(ricaCarrots)
```

Format

'ricaCarrots'

A List of 3 objects:

Y a vector. The production of carrots (open field) (in quintals) for 598 French farms.

X a matrix. The 259 covariates describing the same 598 French farms.

description a matrix. Description of the 259 covariates.

Source

<https://agreste.agriculture.gouv.fr/agreste-web/servicon/I.2/listeTypeServicon/>

SEPaLS — *Function to estimate SEPaLS estimators*

Description

Function to estimate SEPaLS estimators

Usage

```
SEPaLS(
  X,
  Y,
  yn,
  type = c("vMF", "Laplace"),
  mu0 = NULL,
  kappa0 = NULL,
  lambda = NULL
)
```

Arguments

X	$(n \times p)$-dimensional matrix of the covariates.
Y	(n)-dimensional vector of the response.
yn	y_n the quantile correponding to lowest values of Ys to put in th e tail.
type	character, wether vMF for von Mises-Fisher prior or Laplace for Laplace prior. See details.
mu0	μ_0, unitary (p)-dimensional vector. The direction parameter for the vMF prior.
kappa0	κ_0, positive. The concentration parameter for the vMF prior.
lambda	λ, positive. The concentration parameter for the Laplace prior.

Details

The SEPaLS estimators are built depending on the value given to type:

- vMF: then the estimator is proportional to

$$\hat{\beta}_{ml}(y_n) + \kappa_0 \mu_0,$$

where $\hat{\beta}_{ml}(y_n)$ is the EPLS estimator, which coincides with the maximum-likelihood estimator of SEPaLS for a threshold y_n.

- Laplace: then the estimator is proportional to

$$S_\lambda \left(\hat{\beta}_{ml}(y_n) \right),$$

where S_λ is the soft-thresholding operator of threshold λ.

Value

A SEPaLS estimator

See Also

bootstrap.SEPaLS

Examples

```
set.seed(1)
n <- 3000
p <- 10
X <- matrix(rnorm(n*p),n,p)
beta <- c(5:1,rep(0,p-5)) ; beta <- beta/sqrt(sum(beta^2))
Y <- (X%*%beta)^3 + rnorm(n,sd=1/3)
mu0 <- rnorm(p) ; mu0 <- mu0/sqrt(sum(mu0^2))
sepals_vMF <- SEPaLS(X,Y,yn=1,type="vMF",mu0=mu0,kappa0=1)
sepals_Laplace <- SEPaLS(X,Y,yn=1,type="Laplace",lambda=0.01)
```

Index