Package ‘rsig’

February 20, 2015

Type Package
Title Robust Signature Selection for Survival Outcomes
Version 1.0
Author Sangkyun Lee <sangkyun.lee@tu-dortmund.de>, Michel Lang <michellang@gmail.com>
Maintainer Sangkyun Lee <sangkyun.lee@tu-dortmund.de>
Description Robust and efficient feature selection algorithm to identify important features for predicting survival risk. The method is based on subsampling and averaging linear models obtained from the (preconditioned) Lasso algorithm, with an extra shrinking procedure to reduce the size of signatures. An evaluation procedure using subsampling is also provided.
License GPL-2
Depends R (>= 2.15.0), survival, parallel
Imports BBmisc, glmnet, superpc, survcomp, Matrix
Suggests testthat
LazyData yes
ByteCompile yes
NeedsCompilation no
Repository CRAN
Date/Publication 2013-10-12 23:54:18

R topics documented:

 predict.rsig ................................................................. 2
 rsig ................................................................. 2
 rsig.all ................................................................. 4
 rsig.eval ................................................................. 5
Index 7
predict.rsig  

Make Prediction

Description

Return risk prediction on new data.

Usage

```r
## S3 method for class 'rsig'
predict(object, newdata, ...)
```

Arguments

- **object**
  - [rsig]
  - An output object from rsig, see rsig.
- **newdata**
  - [data.frame]
  - Data frame or matrix of input data (rows: examples, columns: features).
- **...**
  - [ANY]
  - Additional arguments, currently ignored.

Value

Risk prediction on new data.

See Also

rsig, rsig.eval, rsig.all

rsig  

Robust Signature Selection for Survival Outcomes

Description

Find a robust signature, i.e. a set of features, using averaged and shrunk generalized linear models. Subsamples are taken to fit models, via $\ell_1$-penalized Cox regression (lasso) or preconditioned lasso (prlasso) algorithm.

Usage

```r
rsig(surv, X, model, n.rep = 10L, plapply = mclapply,
     sd.filter = NULL, verbose = TRUE)
```
**Arguments**

- **surv** [Surv]  
  Survival object, see `Surv`.  
- **X** [data.frame]  
  Data frame or matrix or matrix of input data (rows: examples, columns: features). Columns must have names assigned.  
- **model** [character(1)]  
  Model to use. One of  
  "rs.prlasso" (preconditioned lasso with robust selection),  
  "rs.lasso" (penalized Cox regression with robust selection),  
  "prlasso" (preconditioned lasso), or  
  "lasso" (penalized Cox regression)  
- **n.rep** [integer]  
  The number in replicates to be used for model aggregation. A large enough number is suggested.  
- **plapply** [function]  
  Function used for internal parallelization. Default is `mclapply` for multi-core parallel execution. Change it to `lapply` for single-core execution.  
- **sd.filter** [list]  
  Pre-filter features by their standard deviation, by one of the options specified:  
  `topk`: no. of features to be selected with largest standard deviations, or  
  `quant`: the min percentile in standard deviations of features to be selected.  
- **verbose** [logical]  
  Controls message output.  

**Value**

Object of class “rsig”; a list consisting of  
- **model** model specified by the user  
- **sd.filter** sd.filter object  
- **beta** coefficient vector  
- **intercept** intercept

**See Also**

- `predict.rsig`, `rsig.eval`, `rsig.all`

**Examples**

```r
# An example adapted from glmnet package
set.seed(11011)
n = 300
p = 10
nz = 3
X = matrix(rnorm(n*p),n,p,dimnames=list(NULL,seq_len(p)))
```

```r
beta = rnorm(nz)
f = X[, seq_len(nz)]  # beta
h = exp(f) / 365.25
t = rexp(n, h)
tcens = rbinom(n=n, prob=0.3, size=1)  # censoring indicator
S = Surv(t, 1-tcens)

fit = rsig(S, X, "rs.prlasso", n.rep=2)
pred = predict(fit, X)
perf = rsig.eval(pred, S, X)
```

---

### rsig.all

**Robust Signature Selection for Survival Outcomes with Estimation of Selection Probabilities of Features**

**Description**

Fit a specified model using subsamples and evaluate its performance on out-of-subsample data.

**Usage**

```
rsig.all(surv, X, model, n.rep.out = 10L, n.rep.in = 10L,
         plapply = mclapply, sd.filter = NULL)
```

**Arguments**

- **surv**
  - [Surv]
  - Survival object, see `Surv`.

- **X**
  - [data.frame]
  - Data frame or matrix or matrix of input data (rows: examples, columns: features).

- **model**
  - [character(1)]
  - Model to use. One of "rs.prlasso" (preconditioned lasso with robust selection), "rs.lasso" (penalized Cox regression with robust selection), "prlasso" (preconditioned lasso), or "lasso" (penalized Cox regression).

- **n.rep.out**
  - [integer]
  - The number of replicates to be used to estimate selection probability of features (outer subsampling).

- **n.rep.in**
  - [integer]
  - The number of replicates to be used for model aggregation (inner subsampling).

- **plapply**
  - [function]
  - Function used for internal parallelization. Default is `mclapply` for multi-core parallel execution.
sd.filter [list]
Pre-filter features by their standard deviation, by one of the options specified:
topk: no. of features to be selected with largest standard deviations.
quant: the min percentile in standard deviations of features to be selected.

Value
Object of class “list”.

selection.frequency
a named vector of selected features with their estimated selection frequencies
amongst n.rep.out replicates.

perf
performance measured on out-of-sample data in n.rep.out replicates

See Also
rsig

rsig.eval

Description
Evaluate performance on new data using predictions.

Usage
rsig.eval(pred, surv.new, X.new, measures = "all",
roc.time = 5)

Arguments
pred [predict.rsig]
An output object from predict.rsig, see predict.rsig.
surv.new [Surv]
Survival object, see Surv.
X.new [data.frame]
Data frame or matrix or matrix of input data (rows: examples, columns: fea-
tures).
measures [list]
List of performance measures to be evaluated, "all" or in c("cindex", "tauc")
roc.time [numeric(1)]
Time to evaluate the time-dependent AUC. Default is 5.

Value
Performance values
See Also

rsig.predict.rsig.rsig.all
Index

lapply, 3
mclapply, 3, 4
predict.rsig, 2, 3, 5, 6
rsig, 2, 2, 5, 6
rsig.all, 2, 3, 4, 6
rsig.eval, 2, 3, 5
Surv, 3–5