

Package ‘cornet’

March 21, 2019

Version 0.0.1

Title Elastic Net with Dichotomised Outcomes

Description Implements lasso and ridge regression for dichotomised outcomes (Rauschenberger et al. 2019). Such outcomes are not naturally but artificially binary. They indicate whether an underlying measurement is greater than a threshold.

Depends R (>= 3.0.0)

Imports glmnet, palasso

Suggests knitr, testthat

Enhances RColorBrewer

VignetteBuilder knitr

License GPL-3

LazyData true

Language en-GB

RoxygenNote 6.1.1

URL <https://github.com/rauschenberger/cornet>

BugReports <https://github.com/rauschenberger/cornet/issues>

NeedsCompilation no

Author Armin Rauschenberger [aut, cre]

Maintainer Armin Rauschenberger <a.rauschenberger@vumc.nl>

Repository CRAN

Date/Publication 2019-03-21 17:14:20 UTC

R topics documented:

.check	2
.compare	3
.equal	4
.simulate	4
.test	5

coef.cornet	5
cornet	6
plot.cornet	8
predict.cornet	8
print.cornet	9

Index	11
--------------	-----------

<i>.check</i>	<i>Arguments</i>
---------------	------------------

Description

Verifies whether an argument matches formal requirements.

Usage

```
.check(x, type, dim = NULL, miss = FALSE, min = NULL, max = NULL,
       values = NULL, inf = FALSE, null = FALSE)
```

Arguments

x	argument
type	character "string", "scalar", "vector", "matrix"
dim	vector/matrix dimensionality: integer scalar/vector
miss	accept missing values: logical
min	lower limit: numeric
max	upper limit: numeric
values	only accept specific values: vector
inf	accept infinite (Inf or -Inf) values: logical
null	accept NULL: logical

Examples

```
cornet:::.check(0.5, type="scalar", min=0, max=1)
```

Description

Compares models for a continuous response with a cut-off value.

Usage

```
.compare(y, cutoff, X, alpha = 1, nfolds = 5, foldid = NULL,  
         type.measure = "deviance")
```

Arguments

y	continuous outcome: vector of length n
cutoff	cut-off point for dichotomising outcome into classes: <i>meaningful</i> value between $\min(y)$ and $\max(y)$
X	features: numeric matrix with n rows (samples) and p columns (variables)
alpha	elastic net mixing parameter: numeric between 0 (ridge) and 1 (lasso)
nfolds	number of folds: integer between 3 and n
foldid	fold identifiers: vector with entries between 1 and $nfolds$; or NULL (balance)
type.measure	loss function for binary classification: character "deviance", "mse", "mae", or "class" (see cv.glmnet)

Details

Uses k-fold cross-validation, fits linear, logistic, and combined regression, calculates different loss functions, and examines squared deviance residuals.

Examples

```
n <- 100; p <- 200  
y <- rnorm(n)  
X <- matrix(rnorm(n*p), nrow=n, ncol=p)  
loss <- cornet:::compare(y=y, cutoff=0, X=X)  
loss
```

`.equal` *Equality*

Description

Verifies whether two or more arguments are identical.

Usage

```
.equal(..., na.rm = FALSE)
```

Arguments

<code>...</code>	scalars, vectors, or matrices of equal dimensions
<code>na.rm</code>	remove missing values: logical

Examples

```
cornet:::.equal(1,1,1)
```

`.simulate` *Data simulation*

Description

Simulates data for unit tests

Usage

```
.simulate(n, p, prob = 0.2, fac = 1)
```

Arguments

<code>n</code>	sample size: positive integer
<code>p</code>	covariate space: positive integer
<code>prob</code>	(approximate) proportion of causal covariates: numeric between 0 and 1
<code>fac</code>	noise factor: positive real number

Value

Returns invisible list with elements `y` and `X`.

Examples

```
data <- cornet:::.simulate(n=10,p=20,prob=0.2,fac=2)
names(data)
```

.test *Single-split test*

Description

Compares models for a continuous response with a cut-off value.

Usage

```
.test(y, cutoff, X, alpha = 1, type.measure = "deviance")
```

Arguments

- y continuous outcome: vector of length n
- cutoff cut-off point for dichotomising outcome into classes: *meaningful* value between $\min(y)$ and $\max(y)$
- X features: numeric matrix with n rows (samples) and p columns (variables)
- alpha elastic net mixing parameter: numeric between 0 (ridge) and 1 (lasso)
- type.measure loss function for binary classification: character "deviance", "mse", "mae", or "class" (see [cv.glmnet](#))

Details

Splits samples into 80 calculates squared deviance residuals of logistic and combined regression, conducts the paired one-sided Wilcoxon signed rank test, and returns the p-value.

Examples

```
n <- 100; p <- 200
y <- rnorm(n)
X <- matrix(rnorm(n*p), nrow=n, ncol=p)
cornet::: .test(y=y, cutoff=0, X=X)
```

coef.cornet *Extract estimated coefficients*

Description

Extracts estimated coefficients from linear and logistic regression, under the penalty parameter that minimises the cross-validated loss.

Usage

```
## S3 method for class 'cornet'
coef(object, ...)
```

Arguments

object [cornet](#) object
 ... further arguments (not applicable)

Value

This function returns a matrix with n rows and two columns, where n is the sample size. It includes the estimated coefficients from linear regression (1st column: "beta") and logistic regression (2nd column: "gamma").

Examples

```
n <- 100; p <- 200
y <- rnorm(n)
X <- matrix(rnorm(n*p), nrow=n, ncol=p)
net <- cornet(y=y, cutoff=0, X=X)
coef(net)
```

 cornet

Combined regression

Description

Implements lasso and ridge regression for dichotomised outcomes. Such outcomes are not naturally but artificially binary. They indicate whether an underlying measurement is greater than a threshold.

Usage

```
cornet(y, cutoff, X, alpha = 1, np_i = 101, pi = NULL, nsigma = 99,
       sigma = NULL, n_folds = 10, foldid = NULL,
       type.measure = "deviance", ...)
```

Arguments

y continuous outcome: vector of length n

cutoff cut-off point for dichotomising outcome into classes: *meaningful* value between $\min(y)$ and $\max(y)$

X features: numeric matrix with n rows (samples) and p columns (variables)

alpha elastic net mixing parameter: numeric between 0 (ridge) and 1 (lasso)

np_i number of pi values (weighting)

pi pi sequence: vector of increasing values in the unit interval; or NULL (default sequence)

nsigma number of sigma values (scaling)

sigma sigma sequence: vector of increasing positive values; or NULL (default sequence)

<code>nfolds</code>	number of folds: integer between 3 and n
<code>foldid</code>	fold identifiers: vector with entries between 1 and <code>nfolds</code> ; or NULL (balance)
<code>type.measure</code>	loss function for binary classification: character "deviance", "mse", "mae", or "class" (see <code>cv.glmnet</code>)
<code>...</code>	further arguments passed to <code>glmnet</code>

Details

The argument `family` is unavailable, because this function fits a *gaussian* model for the numeric response, and a *binomial* model for the binary response.

Linear regression uses the loss function "deviance" (or "mse"), but the loss is incomparable between linear and logistic regression.

The loss function "auc" is unavailable for internal cross-validation. If at all, use "auc" for external cross-validation only.

Value

Returns an object of class `cornet`, a list with multiple slots:

- `gaussian`: fitted linear model, class `glmnet`
- `binomial`: fitted logistic model, class `glmnet`
- `sigma`: scaling parameters `sigma`, vector of length `nsigma`
- `pi`: weighting parameters `pi`, vector of length `npi`
- `cvm`: evaluation loss, matrix with `nsigma` rows and `npi` columns
- `sigma.min`: optimal scaling parameter, positive scalar
- `pi.min`: optimal weighting parameter, scalar in unit interval
- `cutoff`: threshold for dichotomisation

References

Armin Rauschenberger and Enrico Glaab (2019). "Lasso and ridge regression for dichotomised outcomes". *Manuscript in preparation*.

See Also

Methods for objects of class `cornet` include `coef` and `predict`.

Examples

```
n <- 100; p <- 200
y <- rnorm(n)
X <- matrix(rnorm(n*p), nrow=n, ncol=p)
net <- cornet(y=y, cutoff=0, X=X)
net
```

`plot.cornet`*Plot loss matrix*

Description

Plots the loss for different combinations of scaling (σ) and weighting (π) parameters.

Usage

```
## S3 method for class 'cornet'  
plot(x, ...)
```

Arguments

`x` [cornet](#) object
`...` further arguments (not applicable)

Value

This function plots the evaluation loss (cvm). Whereas the matrix has σ in the rows, and π in the columns, the plot has σ on the x -axis, and π on the y -axis. For all combinations of σ and π , the colour indicates the loss. If the R package `RColorBrewer` is installed, blue represents low. Otherwise, red represents low. White always represents high.

Examples

```
n <- 100; p <- 200  
y <- rnorm(n)  
X <- matrix(rnorm(n*p), nrow=n, ncol=p)  
net <- cornet(y=y, cutoff=0, X=X)  
plot(net)
```

`predict.cornet`*Predict binary outcome*

Description

Predicts the binary outcome with linear, logistic, and combined regression.

Usage

```
## S3 method for class 'cornet'  
predict(object, newx, type = "probability", ...)
```


Arguments

object	<code>cornet</code> object
newx	covariates: numeric matrix with n rows (samples) and p columns (variables)
type	"probability", "odds", "log-odds"
...	further arguments (not applicable)

Details

For linear regression, this function tentatively transforms the predicted values to predicted probabilities, using a Gaussian distribution with a fixed mean (threshold) and a fixed variance (estimated variance of the numeric outcome).

Examples

```
n <- 100; p <- 200
y <- rnorm(n)
X <- matrix(rnorm(n*p), nrow=n, ncol=p)
net <- cornet(y=y, cutoff=0, X=X)
predict(net, newx=X)
```

print.cornet	<i>Combined regression</i>
--------------	----------------------------

Description

Prints summary of `cornet` object.

Usage

```
## S3 method for class 'cornet'
print(x, ...)
```

Arguments

x	<code>cornet</code> object
...	further arguments (not applicable)

Value

Returns sample size n , number of covariates p , information on dichotomisation, tuned scaling parameter (σ), tuned weighting parameter (π), and corresponding loss.

Examples

```
n <- 100; p <- 200
y <- rnorm(n)
X <- matrix(rnorm(n*p), nrow=n, ncol=p)
net <- cornet(y=y, cutoff=0, X=X)
print(net)
```

Index

.check, 2
.compare, 3
.equal, 4
.simulate, 4
.test, 5

coef, 7
coef.cornet, 5
cornet, 6, 6, 8, 9
cornet-package (cornet), 6
cv.glmnet, 3, 5, 7

glmnet, 7

plot.cornet, 8
predict, 7
predict.cornet, 8
print.cornet, 9