Package 'QuanDA'

November 24, 2025

Title Quantile-Based Discriminant Analysis for High-Dimensional Imbalanced Classification

Version 1.0.0

Description Implements quantile-based discriminant analysis (QuanDA) for imbalanced classification in high-dimensional, low-sample-size settings. The method fits penalized quantile regression directly on discrete class labels and tunes the quantile level to reflect class imbalance.

Depends R (>= 3.5.0)

Imports hdqr, pROC, stats, methods

License GPL-2

NeedsCompilation yes

RoxygenNote 7.2.3

Encoding UTF-8

Author Qian Tang [aut, cre],
Yuwen Gu [aut],
Boxiang Wang [aut]

Maintainer Qian Tang <tang1015@umn.edu>

Repository CRAN

Date/Publication 2025-11-24 09:10:02 UTC

Contents

breast *Example breast cancer data*

Description

A list containing predictor matrix X and binary response y.

Usage

```
data(breast)
```

Value

This data frame contains the following:

x	gene expression levels.
y	Disease state that is coded as 1 and -1

Examples

```
data(breast)
```

predict.quanda *Make Predictions from a ‘quanda‘ Object*

Description

Produces fitted values for new predictor data using a fitted ‘quanda()‘ object.

Usage

```
## S3 method for class 'quanda'
predict(object, newx, type = c("class", "loss"), ...)
```

Arguments

object	Fitted ‘quanda()‘ object from which predictions are to be derived.
newx	Matrix of new predictor values for which predictions are desired. This must be a matrix and is a required argument.
type	Type of prediction required. Type ‘"class"‘ produces the predicted binary class labels and type ‘"loss"‘ returns the fitted values. Default is "class".
...	Not used.

Value

Numeric vector of length n_new.

See Also

quanda

Examples

```
data(breast)
X <- as.matrix(X)
y <- as.numeric(as.character(y))
y[y==-1]=0
fit <- quanda(X, y)
```

---

quanda *Fit QuanDA for imbalanced binary classification*

Description

QuanDA fits a quantile-regression-based discriminant with label jittering. For each candidate quantile level τ, the binary labels are jittered (adding $U(0,1)$), a penalized quantile regression is fit multiple times, and the coefficient vectors are averaged. The best τ is selected by AUC.

Usage

```
quanda(
  x,
  y,
  lambda = 10^(seq(1, -4, length.out = 30)),
  lam2 = 0.01,
  n_rep = 10,
  tau_window = 0.05,
  nfolds = 5,
  maxit = 10000,
  eps = 1e-07,
  maxit_cv = 10000,
  eps_cv = 1e-05
)
```

Arguments

x	A numeric matrix of predictors with n rows (observations) and p columns (features).
y	A binary response vector of length n with values 0 or 1.
lambda	Optional numeric vector of penalty values (largest lambda[1]). If NULL, a default sequence will be generated from the data.
lam2	Numeric, secondary penalty (ridge/elastic term) passed to hdqr. Default 0.01.
n_rep	Integer, number of jittering repetitions (averaged). Default 10.

`tau_window` Width around the class rate to explore quantiles. Candidate τ are $b+\{-w,\ldots,w\}$ in steps of 0.01, clipped to $[0,1]$, where b is the class rate and w is `tau_window`. Default `0.1`.

`nfolds` Integer, number of CV folds used by `cv_z()`. Default `5`.

`maxit`, `maxit_cv`, `eps`, `eps_cv`
Controls for inner optimizers and CV helper.

Details

We jitter labels via $z_i = y_i + U_i$, where $U_i \sim \mathrm{Unif}(0,1)$, fit penalized quantile regression at multiple τ, average coefficients over `n_rep` jitters, compute AUCs on the original (x, y), and pick the τ that maximizes AUC.

Value

An object of class `"quanda"` with elements:

beta Numeric vector of length $p+1$ (intercept first).

tau_grid Numeric vector of candidate τ values.

tau_best Chosen τ.

auc Vector of AUCs across τ.

call The matched call.

Examples

```
data(breast)
X <- as.matrix(X)
y <- as.numeric(as.character(y))
y[y==-1]=0
fit <- quanda(X, y)
pred <- predict(fit, tail(X))
```

Index