Package 'regda'

May 9, 2026

Type Package

Title Regularised Discriminant Analysis

Version 1.0

Date 2023-11-05

Author Michail Tsagris [aut, cre]

Maintainer Michail Tsagris <mtsagris@uoc.gr>

Depends R (>= 4.0)

Imports doParallel, foreach, parallel, Rfast, Rfast2, stats

Description

Regularised discriminant analysis functions. The classical regularised discriminant analysis proposed by Friedman in 1989, including cross-validation, of which the linear and quadratic discriminant analyses are special cases. Further, the regularised maximum likelihood linear discriminant analysis, including cross-validation. References: Friedman J.H. (1989). ``Regularized Discriminant Analysis". Journal of the American Statistical Association 84(405): 165--175. <doi:10.2307/2289860>. Friedman J., Hastie T. and Tibshirani R. (2009). ``The elements of statistical learning", 2nd edition. Springer, Berlin. <doi:10.1007/978-0-387-84858-7>. Tsagris M., Preston S. and Wood A.T.A. (2016). ``Improved classification for compositional data using the alpha-transformation". Journal of Classification, 33(2): 243--261. <doi:10.1007/s00357-016-9207-5>.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2023-11-06 17:40:02 UTC

Contents

regda-package *Regularised Discriminant Analysis*

Description

Description: Regularised discriminant analysis functions. The classical regularised discriminant analysis proposed by Friedman in 1989, including cross-validation, of which the linear and quadratic discriminant analyses are special cases. Further, the regularised maximum likelihood linear discriminant analysis, including cross-validation.

Details

Package:	regda
Type:	Package
Version:	1.0
Date:	2023-11-05
License:	GPL-2

Maintainers

Michail Tsagris <mtsagris@uoc.gr>.

Author(s)

Michail Tsagris <mtsagris@uoc.gr>

References

Friedman J.H. (1989): Regularized Discriminant Analysis. Journal of the American Statistical Association 84(405): 165–175.

Friedman Jerome, Trevor Hastie and Robert Tibshirani (2009). The elements of statistical learning, 2nd edition. Springer, Berlin.

Tsagris M., Preston S. and Wood A.T.A. (2016). Improved classification for compositional data using the α-transformation. Journal of Classification, 33(2): 243–261.

`Cross-validation for the regularised maximum likelihood linear discriminant analysis`

Cross-validation for the regularised maximum likelihood linear discriminant analysis

Description

Cross-validation for the regularised maximum likelihood linear discriminant analysis.

Usage

```
regmlelda.cv(x, ina, lambda = seq(0, 1, by = 0.1), folds = NULL, nfolds = 10,
             stratified = TRUE, seed = FALSE, pred.ret = FALSE)
```

Arguments

`x`	A matrix with numerical data.
`ina`	A numerical vector or factor with consecutive numbers indicating the group to which each observation belongs to.
`lambda`	A vector of regularization values λ such as (0, 0.1, 0.2,...).
`folds`	A list with the indices of the folds.
`nfolds`	The number of folds to be used. This is taken into consideration only if "folds" is NULL.
`stratified`	Do you want the folds to be selected using stratified random sampling? This preserves the analogy of the samples of each group. Make this TRUE if you wish.
`seed`	If you set this to TRUE, the same folds will be created every time.
`pred.ret`	If you want the predicted values returned set this to TRUE.

Details

Cross-validation for the regularised maximum likelihood linear discriminant analysis is performed. The function is not extremely fast, yet is pretty fast.

Value

A list including:

`preds`	If pred.ret is TRUE the predicted values for each fold are returned as elements in a list.
`crit`	A vector whose length is equal to the number of k and is the accuracy metric for each k. For the classification case it is the percentage of correct classification. For the regression case the mean square of prediction error.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Friedman J., Hastie T. and Tibshirani R. (2017). The elements of statistical learning. New York: Springer.

See Also

reg.mle.lda

Examples

```
x <- as.matrix(iris[, 1:4])
mod <- regmlelda.cv(x, iris[, 5])
```

Regularised discriminant analysis for Euclidean data

Regularised discriminant analysis for Euclidean data

Description

Regularised discriminant analysis for Euclidean data.

Usage

```
rda(xnew, x, ina, gam = 1, del = 0)
```

Arguments

xnew	A matrix with the new data whose group is to be predicted. They have to be continuous.
x	A matrix with the available data. They have to be continuous.
ina	A group indicator variable for the avaiable data.
gam	This is a number between 0 and 1. It is the weight of the pooled covariance and the diagonal matrix.
del	This is a number between 0 and 1. It is the weight of the LDA and QDA.

Details

The covariance matrix of each group is calculated and then the pooled covariance matrix. The spherical covariance matrix consists of the average of the pooled variances in its diagonal and zeros in the off-diagonal elements. gam is the weight of the pooled covariance matrix and 1-gam is the weight of the spherical covariance matrix, Sa = gam * Sp + (1-gam) * sp. Then it is a compromise between LDA and QDA. del is the weight of Sa and 1-del the weight of each group covariance group.

Value

A list including:

`prob`	The estimated probabilities of the new data of belonging to each group.
`scores`	The estimated socres of the new data of each group.
`est`	The estimated group membership of the new data.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Friedman J.H. (1989): Regularized Discriminant Analysis. Journal of the American Statistical Association 84(405): 165–175.

Friedman Jerome, Trevor Hastie and Robert Tibshirani (2009). The elements of statistical learning, 2nd edition. Springer, Berlin.

Tsagris M., Preston S. and Wood A.T.A. (2016). Improved classification for compositional data using the α-transformation. Journal of Classification, 33(2): 243–261.

See Also

`rda.tune`

Examples

```
x <- as.matrix(iris[, 1:4])
ina <- iris[, 5]
mod <- rda(x, x, ina)
table(ina, mod$est)
```

`Regularised maximum likelihood linear discriminant analysis`

Regularised maximum likelihood linear discriminant analysis

Description

Regularised maximum likelihood linear discriminant analysis.

Usage

```
reg.mle.lda(xnew, x, ina, lambda)
```

Arguments

xnew	A numerical vector or a matrix with the new observations, continuous data.
x	A matrix with numerical data.
ina	A numerical vector or factor with consecutive numbers indicating the group to which each observation belongs to.
lambda	A vector of regularization values λ such as (0, 0.1, 0.2,...).

Details

Regularised maximum likelihood linear discriminant analysis is performed. The function is not extremely fast, yet is pretty fast.

Value

A matrix with the predicted group of each observation in "xnew". Every column corresponds to a λ value. If you have just on value of λ, then you will have one column only.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

See Also

regmlelda.cv

Examples

```
x <- as.matrix(iris[, 1:4])
ina <- iris[, 5]
a <- reg.mle.lda(x, x, ina, lambda = seq(0, 1, by = 0.1) )
```

Tuning the parameters of the regularised discriminant analysis

Tuning the parameters of the regularised discriminant analysis

---

Description

Tuning the parameters of the regularised discriminant analysis for Eucldiean data.

Usage

```
rda.tune(x, ina, nfolds = 10, gam = seq(0, 1, by = 0.1), del = seq(0, 1, by = 0.1),
ncores = 1, folds = NULL, stratified = TRUE, seed = NULL)
```

Arguments

x	A matrix with the data.
ina	A group indicator variable for the avaiable data.
nfolds	The number of folds in the cross validation.
gam	A grid of values for the γ parameter as defined in Tsagris et al. (2016).
del	A grid of values for the δ parameter as defined in Tsagris et al. (2016).
ncores	The number of cores to use. If more than 1, parallel computing will take place. It is advisable to use it if you have many observations and or many variables, otherwise it will slow down th process.
folds	If you have the list with the folds supply it here. You can also leave it NULL and it will create folds.
stratified	Do you want the folds to be created in a stratified way? TRUE or FALSE.
seed	You can specify your own seed number here or leave it NULL.

Details

Cross validation is performed to select the optimal parameters for the regularisded discriminant analysis and also estimate the rate of accuracy.

The covariance matrix of each group is calcualted and then the pooled covariance matrix. The spherical covariance matrix consists of the average of the pooled variances in its diagonal and zeros in the off-diagonal elements. gam is the weight of the pooled covariance matrix and 1-gam is the weight of the spherical covariance matrix, Sa = gam * Sp + (1-gam) * sp. Then it is a compromise between LDA and QDA. del is the weight of Sa and 1-del the weight of each group covariance group.

Value

A list including: If graph is TRUE a plot of a heatmap of the performance s will appear.

per	An array with the estimate rate of correct classification for every fold. For each of the M matrices, the row values correspond to gam and the columns to the del parameter.
percent	A matrix with the mean estimated rates of correct classification. The row values correspond to gam and the columns to the del parameter.
se	A matrix with the standard error of the mean estimated rates of correct classification. The row values correspond to gam and the columns to the del parameter.
result	The estimated rate of correct classification along with the best gam and del parameters.
runtime	The time required by the cross-validation procedure.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Friedman J.H. (1989): Regularized Discriminant Analysis. Journal of the American Statistical Association 84(405): 165–175.

Friedman Jerome, Trevor Hastie and Robert Tibshirani (2009). The elements of statistical learning, 2nd edition. Springer, Berlin.

Tsagris M., Preston S. and Wood A.T.A. (2016). Improved classification for compositional data using the α-transformation. Journal of Classification, 33(2): 243–261.

See Also

rda

Examples

```
mod <- rda.tune(as.matrix(iris[, 1:4]), iris[, 5], gam = seq(0, 1, by = 0.2),
del = seq(0, 1, by = 0.2) )
mod
```

Index