Package 'lincom'

July 1, 2024

Version 1.2

Date 2024-06-30

Title Linear Biomarker Combination: Empirical Performance Optimization

Author Yijian Huang <yhuang5@emory.edu>

Maintainer Yijian Huang <yhuang5@emory.edu>

Depends R (>= 3.6.0)

Imports SparseM, Rmosek, methods, stats

SystemRequirements MOSEK (>= 6), MOSEK License (>= 6)

Suggests knitr, rmarkdown

VignetteBuilder knitr

Description Perform two linear combination methods for biomarkers: (1) Empirical performance optimization for specificity (or sensitivity) at a controlled sensitivity (or specificity) level of Huang and Sanda (2022) <doi:10.1214/22-aos2210>, and (2) weighted maximum score estimator with empirical minimization of averaged false positive rate and false negative rate. Both adopt the algorithms of Huang and Sanda (2022) <doi:10.1214/22-aos2210>. 'MOSEK' solver is used and needs to be installed; an academic license for 'MOSEK' is free.

License GPL (>= 2)

NeedsCompilation yes

Repository CRAN

Date/Publication 2024-06-30 23:30:02 UTC

Contents

| eum | *Empirical performance optimization for specificity (or sensitivity) at a controlled sensitivity (or specificity) level* |

Description

Linear combination of multiple biomarkers

Usage

```
eum(mk, n1, s0, w=2, grdpt=10, contract=0.8, fixsens=TRUE, lbmdis=TRUE)
```

Arguments

mk	biomarker values of cases followed by controls, with each row containing multiple markers from an individual.
n1	size of cases.
s0	controlled level of sensitivity or specificity.
w	weight for l1 norm of combination coefficient in the objective function (w>1 guarantees sound asymptotic properties).
grdpt	number of grid points in coarse grid search for initial value; if grdpt=0, use logistic regression instead.
contract	reduction factor in the sequence of approximation parameters for indicator function.
fixsens	fixing sensitivity if True, and specificity otherwise.
lbmdis	larger biomarker value is more associated with cases if True, and controls otherwise.

Value

coef	estimated combination coefficient, with unity l1 norm.
hs	empirical estimate of specificity at controlled sensitivity, or vice versa.
threshold	estimated threshold.
init_coef	initial combination coefficient, with unity l1 norm.
init_hs	initial specificity at controlled sensitivity, or vice versa.
init_threshold	estimated threshold for the initial combination coefficient.

Author(s)

Yijian Huang

References

Huang and Sanda (2022). Linear biomarker combination for constrained classification. *The Annals of Statistics* 50, 2793–2815

Examples

```
## simulate 3 biomarkers for 100 cases and 100 controls
mk <- rbind(matrix(rnorm(300),ncol=3),matrix(rnorm(300),ncol=3))
mk[1:100,1] <- mk[1:100,1]/sqrt(2)+1
mk[1:100,2] <- mk[1:100,2]*sqrt(2)+1

## linear combination to empirically maximize specificity at controlled 0.95
## sensitivity
## Require installation of 'MOSEK' to run
## Not run:
lcom <- eum(mk, 100, 0.95, grdpt=0)

## End(Not run)
```

---

wmse	*Weighted Manski's maximum score estimator*

Description

empirical minimization of averaged false positive rate and false negative rate

Usage

```
wmse(mk, n1, r=1, w=2, contract=0.8, lbmdis=TRUE)
```

Arguments

mk	biomarker values of cases followed by controls, with each row containing multiple markers from an individual.
n1	size of cases.
r	weight of false positive rate relative to false negative rate.
w	weight for l1 norm of combination coefficient in the objective function (w>1 guarantees sound asymptotic properties).
contract	reduction factor in the sequence of approximation parameters for indicator function.
lbmdis	larger biomarker value is more associated with cases if True, and controls otherwise.

Value

coef	estimated combination coefficient, with unity l1 norm.
obj	empirical objective function: r * false positive rate + false negative rate.
threshold	estimated threshold.
init_coef	initial combination coefficient from logistic regression, with unity l1 norm.
init_obj	empirical objective function for the initial combination coefficient.
init_threshold	estimated threshold for the initial combination coefficient.

Author(s)

Yijian Huang

References

Huang and Sanda (2022). Linear biomarker combination for constrained classification. *The Annals of Statistics* 50, 2793–2815

Examples

```
## simulate 3 biomarkers for 100 cases and 100 controls
mk <- rbind(matrix(rnorm(300),ncol=3),matrix(rnorm(300),ncol=3))
mk[1:100,1] <- mk[1:100,1]/sqrt(2)+1
mk[1:100,2] <- mk[1:100,2]*sqrt(2)+1

## linear combination to empirically minimize averaged false positive rate and
## false negative rate
## Require installation of 'MOSEK' to run
## Not run:
lcom <- wmse(mk, 100)

## End(Not run)
```

Index