

Package ‘FARDEEP’

January 20, 2025

Title Fast and Robust Deconvolution of Tumor Infiltrating Lymphocyte from Expression Profiles using Least Trimmed Squares

Version 1.0.1

Description Using the idea of least trimmed square, it could automatically detects and removes outliers from data before estimating the coefficients. It is a robust machine learning tool which can be applied to gene-expression deconvolution technique. Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie (2019) <[doi:10.1101/358366](https://doi.org/10.1101/358366)>.

Depends R (>= 3.3.0)

License MIT + file LICENSE

Imports nnls(>= 1.4), stats, preprocessCore

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

NeedsCompilation no

Author Yuning Hao [aut],
Ming Yan [aut],
Blake R. Heath [aut],
Yu L. Lei [aut],
Yuying Xie [aut, cre]

Maintainer Yuying Xie <xyy@egr.msu.edu>

Repository CRAN

Date/Publication 2019-04-24 15:10:03 UTC

Contents

alts	2
fardeep	3
LM22	4
mixture	5
sample.sim	6
tuningBIC	7

Index	9
--------------	----------

alts *Using the basic idea of least trimmed square to detect and remove outliers before estimating the coefficients. Adaptive least trimmed square.*

Description

Using the basic idea of least trimmed square to detect and remove outliers before estimating the coefficients. Adaptive least trimmed square.

Usage

```
alts(x, y, alpha1 = 0.1, alpha2 = 1.5, k = 6, nn = TRUE,
     intercept = TRUE)
```

Arguments

x	input matrix of predictors with n rows and p columns.
y	input vector of dependent variable with length n.
alpha1	parameter used to adjust the upper bound of outliers. Take value from 0 to 1, default 0.1.
alpha2	parameter used to adjust the lower bound of outliers. Take value larger than 1, default 1.5.
k	parameter used to determine the boundary of outliers in the following step of algorithm. Take value from 1 to 10, default 6.
nn	whether coefficients are non-negative, default TRUE.
intercept	whether intercept is included in model, default TRUE.

Value

beta: estimation of coefficients.
 number_outlier: number of outliers.
 outlier_detect: index of detected outliers.
 X.new: good observed points for independent variables.
 Y.new: good observed points for dependent variables.
 k: modified k (if the input value is not appropriate).

Author(s)

Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie

References

Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie. Fast and Robust Deconvolution of Tumor Infiltrating Lymphocyte from Expression Profiles using Least Trimmed Squares. <doi:10.1101/358366>

Examples

```
library(FARDEEP)
samp = sample.sim(n = 500, p = 20, sig = 1, a1 = 0.1, a2 = 0.2, nn = TRUE, intercept = TRUE)
result = alts(samp$x, samp$y, alpha1 = 0.1, alpha2 = 1.5, k = 6, nn = TRUE, intercept = TRUE)
coef = result$beta
```

fardeep	<i>Using the idea of least trimmed square to detect and remove outliers before estimating the coefficients. A robust method for gene-expression deconvolution.</i>
---------	--

Description

Using the idea of least trimmed square to detect and remove outliers before estimating the coefficients. A robust method for gene-expression deconvolution.

Usage

```
fardeep(X, Y, alpha1 = 0.1, alpha2 = 1.5, up = 10, low = 1,
        nn = TRUE, intercept = TRUE, lognorm = TRUE, permn = 100,
        QN = FALSE)
```

Arguments

X	input matrix of predictors with n rows and p columns.
Y	input vector of dependent variable.
alpha1	parameter used to adjust the upper bound of outliers. Take value from 0 to 1, default 0.1.
alpha2	parameter used to adjust the lower bound of outliers. Take value larger than 1, default 1.5.
up	upper bound of parameter k in function alts, default 10.
low	lower bound of parameter k in function alts, default 1.
nn	whether coefficients are non-negative, default TRUE.
intercept	whether intercept is included in model, default TRUE.
lognorm	whether noise is log-normal distributed, default TRUE.
permn	the number of permutation to get the p-values, default TRUE.
QN	whether perform quantile normalization, default TRUE.

Value

abs.beta: estimation of absolute abundance of cells (TIL subset scores).
 relative.beta: estimation of relative proportions by normalizing abs.beta to 1.
 pval: statistical significance for the deconvolution result.
 k.value: tuned parameter by modified BIC.

Author(s)

Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie

References

Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie. Fast and Robust Deconvolution of Tumor Infiltrating Lymphocyte from Expression Profiles using Least Trimmed Squares. <doi:10.1101/358366>

Examples

```
library(FARDEEP)
data(LM22)
data(mixture)
# toy examples
result = fardeep(LM22, mixture[, 1:2], permn = 0)

result = fardeep(LM22, mixture)
coef = result$abs.beta
```

LM22

Signature matrix

Description

A dataset containing 547 genes and 22 TILs.

Usage

LM22

Format

A data frame with 547 rows and 22 variables:

B.cells.naive naive B cells

B.cells.memory memory B cells

Plasma.cells Plasma cells

T.cells.CD8 CD8 T cells

T.cells.CD4.naive naive CD4 T cells

T.cells.CD4.memory.resting resting memory CD4 T cells

T.cells.CD4.memory.activated activated memory CD4 T cells

T.cells.follicular.helper follicular helper T cells

T.cells.regulatory.Tregs. regulatory T cells

T.cells.gamma.delta gamma delta T cells
NK.cells.resting resting natural killer cells
NK.cells.activated activated natural killer cells
Monocytes monocytes
Macrophages.M0 M0 macrophages
Macrophages.M1 M1 macrophages
Macrophages.M2 M2 macrophages
Dendritic.cells.resting resting dendritic cells
Dendritic.cells.activated activated dendritic cells
Mast.cells.resting resting mast cells
Mast.cells.activated activated mast cells
Eosinophils eosinophils
Neutrophils neutrophils

References

Aaron M. Newman, Chih Long Liu, Michael R. Green, Andrew J. Gentles, Weiguo Feng, Yue Xu, Chuong D. Hoang, Maximilian Diehn and Ash A. Alizadeh. Robust enumeration of cell subsets from tissue expression profiles.

mixture

Gene-expression data from 14 follicular lymphoma patients

Description

This gene-expression dataset consists of surgical lymph node biopsies of 14 follicular lymphoma patients with 19416 genes. It is available on Gene Expression Omnibus (GEO) with accession number GSE65135. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65135>.

Usage

mixture

Format

A data frame with 19416 rows and 14 variables:

GSM1587831 FL lymph node biopsy, untreated, 1063
GSM1587832 FL lymph node biopsy, untreated, 1080
GSM1587833 FL lymph node biopsy, untreated, 575
GSM1587834 FL lymph node biopsy, untreated, 581
GSM1587835 FL lymph node biopsy, untreated, 598

GSM1587836 FL lymph node biopsy, untreated, 639
GSM1587837 FL lymph node biopsy, untreated, 664
GSM1587838 FL lymph node biopsy, untreated, 666
GSM1587839 FL lymph node biopsy, untreated, 695
GSM1587840 FL lymph node biopsy, untreated, 706
GSM1587841 FL lymph node biopsy, untreated, 726
GSM1587842 FL lymph node biopsy, untreated, 731
GSM1587843 FL lymph node biopsy, untreated, 944
GSM1587844 FL lymph node biopsy, untreated, 959

References

Aaron M. Newman, Chih Long Liu, Michael R. Green, Andrew J. Gentles, Weiguo Feng, Yue Xu, Chuong D. Hoang, Maximilian Diehn and Ash A. Alizadeh. Robust enumeration of cell subsets from tissue expression profiles.

sample.sim	<i>Generate random sample with different proportion of outliers and leverage points</i>
------------	---

Description

Generate random sample with different proportion of outliers and leverage points

Usage

```
sample.sim(n, p, sig, a1, a2, nn = TRUE, intercept = FALSE)
```

Arguments

n	number of observations.
p	number of independent variables (predictors).
sig	variance of dependent variable.
a1	proportion of outliers.
a2	proportion of leverage points in outliers.
nn	whether coefficients are non-negative, default TRUE.
intercept	whether intercept is included in model, default TRUE.

Value

y: vector of dependent variable.
x: matrix of predictors with n rows and p columns.
loc: index of added outliers.
beta: vector of coefficients.

Author(s)

Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie

References

Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie. Fast and Robust Deconvolution of Tumor Infiltrating Lymphocyte from Expression Profiles using Least Trimmed Squares. <doi:10.1101/358366>

Examples

```
library(FARDEEP)
samp = sample.sim(n = 500, p = 20, sig = 1, a1 = 0.1, a2 = 0.2, nn = TRUE, intercept = TRUE)
```

tuningBIC	<i>Tuning parameter k in function alts using Bayesian Information Criterion (BIC) with some adjustment.</i>
-----------	---

Description

Tuning parameter k in function alts using Bayesian Information Criterion (BIC) with some adjustment.

Usage

```
tuningBIC(x, y, alpha1 = 0.1, alpha2 = 1.5, up = 10, low = 1,
          nn = TRUE, intercept = TRUE, lognorm = TRUE)
```

Arguments

x	input matrix of predictors with n rows and p columns.
y	input vector of dependent variable with length n.
alpha1	parameter used to adjust the upper bound of outliers. Take value from 0 to 1, default 0.1.
alpha2	parameter used to adjust the lower bound of outliers. Take value larger than 1, default 1.5.
up	upper bound of parameter k in function alts, default 10.
low	lower bound of parameter k in function alts, default 1.
nn	whether coefficients are non-negative, default TRUE.
intercept	whether intercept is included in model, default TRUE.
lognorm	whether noise is log-normal distributed, default TRUE.

Value

k: tuning result of parameter k for function alts.

Author(s)

Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie

References

Yuning Hao, Ming Yan, Blake R. Heath, Yu L. Lei and Yuying Xie. Fast and Robust Deconvolution of Tumor Infiltrating Lymphocyte from Expression Profiles using Least Trimmed Squares. <doi:10.1101/358366>

Examples

```
library(FARDEEP)
samp = sample.sim(n = 500, p = 20, sig = 1, a1 = 0.1, a2 = 0.2, nn = TRUE, intercept = TRUE)
k = tuningBIC(samp$x, samp$y, lognorm = FALSE)
```


Index

* **datasets**

LM22, 4

mixture, 5

alts, 2

fardeep, 3

LM22, 4

mixture, 5

sample.sim, 6

tuningBIC, 7