Package 'rcppreqtl'

September 12, 2018

Type Package

Title eqtl optimization using 'RcppEigen'

Version 0.99.0

Date 2016-01-19 The 'rcppreqtl' package uses 'RcppEigen' and C++11 numeric solver for eqtl optimization.

Author Vasyl Zhabotynsky [aut, cre], Wei Sun [aut]

Maintainer Vasyl Zhabotynsky <vasyl@unc.edu>

Description
Analysis of combined total and allele specific reads from the human trio experiment using RNA-seq data.

License GPL (>= 2) | file LICENSE

Depends R (>= 2.15.1),MASS,VGAM

LazyLoad yes

LinkingTo RcppEigen, Rcpp

Imports Matrix (>= 1.1-0), RcppEigen (>= 0.3.2.0), Rcpp (>= 0.11.0), stats, utils

NeedsCompilation yes

Archs i386, x64

R topics documented:

`fit` *Optimization wrapper, maximizing the joint model of total (TReC) and allele specific (ASE) counts for autosomes*

Description

Performs optimization of joint TReC and ASE for autosome and tests with lrt test for two hypotheses hypotheses: additive and parent of origin.

Usage

```
fit(subset=NULL, data, traceit=FALSE)
```

Arguments

`subset`	a subset of entries to be tested, but default is set to NULL which leads to fitting all the genes in the table
`data`	an object of class readCounts including read counts and necessary supporting information
`traceit`	include more debug output, by default set to FALSE

Value

a list of following matrices:

`full`	matrix with columns: log(phiNB) and -log(phiBB), coefficients, -log(likelihood) for the full model fit as well as appended two p-value tests for additive and parent of origin effect
`testadd`	matrix with columns: log(phiNB) and -log(phiBB), coefficients, -log(likelihood) for the restricted to b0=0 model fit
`testpoo`	matrix with columns: log(phiNB) and -log(phiBB), coefficients, -log(likelihood) for the restricted to b1=0 model fit

Author(s)

Vasyl Zhabotynsky vasyl@unc.edu

See Also

`fitsh`, `data`, `simu4`, `simu2`, `readCounts`.

Examples

```
## Not run:
# fitting autosome data for a full model with allele-specific counts collected on gene level:
percase = 0.1
dblcnt = 0.2
mn = 100
b0 = 0
b1 = 0
phiNB = .5;
phiBB=phiNB/4
```

```
niter = 10
betas = c(3,.2,.05,.5)
ss=2

dep = makeXmatr(ss)
dat = simu4(num=niter, Xmatr=dep$Xmatr, haplotype=dep$thp, totmean=mn, percase=percase, dblcnt=dblcnt, phiNB=
fit(subset=NULL, data=dat, traceit=FALSE)

## End(Not run)
```

`fitsh` *Optimization wrapper, maximizing the joint model of total (TReC) and allele specific (ASE) counts for autosomes*

Description

Performs optimization of joint TReC and ASE for autosome and tests with lrt test for two hypotheses hypotheses: additive and parent of origin. This function modification assumes that phiNB and phiBB are common for the gene of interest. Otherwise the model is exactly the same.

Usage

```
fitsh(subset=NULL, data, traceit=FALSE)
```

Arguments

`subset`	a subset of entries to be tested, but default is set to NULL which leads to fitting all the genes in the table
`data`	an object of class readCounts including read counts and necessary supporting information
`traceit`	include more debug output, by default set to FALSE

Value

a list of following matrices:

`full`	matrix with columns: log(phiNB) and -log(phiBB), coefficients, -log(likelihood) for the full model fit as well as appended two p-value tests for additive and parent of origin effect
`testadd`	matrix with columns: log(phiNB) and -log(phiBB), coefficients, -log(likelihood) for the restricted to b0=0 model fit
`testpoo`	matrix with columns: log(phiNB) and -log(phiBB), coefficients, -log(likelihood) for the restricted to b1=0 model fit

Author(s)

Vasyl Zhabotynsky vasyl@unc.edu

See Also

`fitsh`, `data`, `simu2`, `simu4`, `readCounts`.

Examples

```
## Not run:
# fitting autosome data for a full model with allele-specific counts collected on gene level:
percase = 0.1
dblcnt = 0.2
mn = 100
b0 = 0
b1 = 0
phiNB = .5;
phiBB=phiNB/4
niter = 10
betas = c(3,.2,.05,.5)
ss=2

dep = makeXmatr(ss)
dat = simu4(num=niter, Xmatr=dep$Xmatr, haplotype=dep$thp, totmean=mn, percase=percase, dblcnt=dblcnt, phiNB=
fitsh(subset=NULL, data=dat, traceit=FALSE)

## End(Not run)
```

`makeXmatr` *Create example design matrix for simulations*

Description

Produces a design matrix of several sample sizes which can be used to generate simulated dataset

Usage

```
makeXmatr(ss)
```

Arguments

`ss`	Sample size class: ss=1 implies dample size 32, ss=2 implies sample size 64, etc

Value

a design matrix of 4 variables

Author(s)

Vasyl Zhabotynsky vasyl@unc.edu

See Also

`fitsh`, `data`, `simu4`, `simu2`, `readCounts`.

Examples

```
## Not run:
# fitting autosome data for a full model with allele-specific counts collected on gene level:
percase = 0.1
dblcnt = 0.2
mn = 100
b0 = 0
b1 = 0
phiNB = .5;
phiBB=phiNB/4
niter = 100
betas = c(3,.2,.05,.5)
ss=2

dep = makeXmatr(ss)

## End(Not run)
```

`readCounts` *A list object that should be used as input to optimization fit or fitsh function.*

Description

It should contain at least total read counts (TReC), overall allele-specific counts (ASE), paternal allele-specific counts and haplotype classification 0 to 4. Also should include X matrix with covariates such as intercept, library depth, principal components.

Value

`haplotype`	a matrix defining the haplotype status of the individual for each gene - AB=0, BA=1, AA=2, BB=3 coded as PaternalMaternal haplotypes. Each row - one gene
`trc`	matrix of TReC counts. Each row - one gene
`asn`	matrix of ASE counts for subject (column) for corresponding genes (row).
`asnp`	matrix of ASE counts belonging to paternal allele, for correponding subjects and genes as in **asn**.
`haplotypeA`	haplotype modification for a setup with allele-specific reads collected for multiple SNPs in a gene. This block by default is NULL, but can be used by simulation function to study scenarios when reads are collected not on gene level, but on SNP level
`asnA`	allele-specific count corresponding to **haplotypeA**
`asnpA`	paternal allele-specific count corresponding to **haplotypeA**
`X`	design matrix for total read counts - a place to include intercept, library depth, other covariates such as batch effect
`params`	if data is produced by simulation function parameter values can be stored here for further comparisons
`settings`	other important settings for the simulation such as mean of total read counts, percentage of allele specific counts, percentage of reads double-counted with neighboring SNPs can be stored here

Author(s)

Vasyl Zhabotynsky vasyl@unc.edu

See Also

fit, fitsh, simu2, simu4.

Examples

```
## Not run:
# see total read counts (TReC) for first 2 X chromosome genes of a data example:
rc = readCounts(haplotype=haplotypef, trc=trc, asn=asnf, asnp=asnpf, haplotypeA=haplotyped, asnA=asnm, asnpA=
        X=Xmatr, params=c(phiNB, phiBB, b0, b1, betas), settings=c(totmean, percase, dblcnt))

## End(Not run)
```

---

simu2 *Simulate a dataset in a format acceptable by a fit function*

Description

Creates an object of a class readCounts with data simulated according to a provided setup

Usage

```
 simu2(num, Xmatr, haplotype, totmean, percase=0.1, dblcnt=0, phiNB=1, phiBB=0.5, b0=0, b1=0, betas
fullest = fit(subset=NULL, data=dat, traceit=FALSE)
```

Arguments

num	number of iterations
Xmatr	design matrix for total read counts covariates
haplotype	classes of haplotypes 0 - AA, 1 - AB, 2 - BA, 3 - BB where first letter represents paternal allele
totmean	average total gene expression
percase	percentage of reads classified as allele-specific, default value 10%
dblcnt	optional output considering a simulation split into 2 SNPs with double-counting. Default value 0.
phiNB	over-dispersion for Negative-Binomial distribution, default value 1
phiBB	over-dispersion for Beta-Binomial distribution, default value 0.5
b0	additive eQTL, default value 0
b1	parent of origin effect, default value 0
betas	covariates for design matrix Xmatr

Value

an object of class readCounts
simulated dataset that can be used to fit the model

Author(s)

Vasyl Zhabotynsky vasyl@unc.edu

See Also

fitsh, data, simu4, simu2, readCounts.

Examples

```
## Not run:
# fitting autosome data for a full model with allele-specific counts collected on gene level:
percase = 0.1
dblcnt = 0.2
mn = 100
b0 = 0
b1 = 0
phiNB = .5;
phiBB=phiNB/4
niter = 100
betas = c(3,.2,.05,.5)
ss=2

dep = makeXmatr(ss)
dat = simu2(num=niter, Xmatr=dep$Xmatr, haplotype=dep$thp, totmean=mn, percase=percase, dblcnt=dblcnt, phiNB=

## End(Not run)
```

---

simu4 *Simulate a dataset in a format acceptable by a fit function*

Description

Creates an object of a class readCounts with data simulated according to a provided setup

Usage

```
 simu4(num, Xmatr, haplotype, totmean, percase=0.1, dblcnt=0, phiNB=1, phiBB=0.5, b0=0, b1=0, betas
fullest = fit(subset=NULL, data=dat, traceit=FALSE)
```

Arguments

num	number of iterations
Xmatr	design matrix for total read counts covariates
haplotype	classes of haplotypes 0 - AA, 1 - AB, 2 - BA, 3 - BB where first letter represents paternal allele
totmean	average total gene expression
percase	percentage of reads classified as allele-specific, default value 10%
dblcnt	optional output considering a simulation split into 4 SNPs with double-counting. Default value 0.
phiNB	over-dispersion for Negative-Binomial distribution, default value 1

phiBB	over-dispersion for Beta-Binomial distribution, default value 0.5
b0	additive eQTL, default value 0
b1	parent of origin effect, default value 0
betas	covariates for design matrix Xmatr

Value

an object of class readCounts
: simulated dataset that can be used to fit the model

Author(s)

Vasyl Zhabotynsky vasyl@unc.edu

See Also

fitsh, data, simu4, simu2, readCounts.

Examples

```
## Not run:
# fitting autosome data for a full model with allele-specific counts collected on gene level:
percase = 0.1
dblcnt = 0.2
mn = 100
b0 = 0
b1 = 0
phiNB = .5;
phiBB=phiNB/4
niter = 100
betas = c(3,.2,.05,.5)
ss=2

dep = makeXmatr(ss)
dat = simu4(num=niter, Xmatr=dep$Xmatr, haplotype=dep$thp, totmean=mn, percase=percase, dblcnt=dblcnt, phiNB=

## End(Not run)
```

Index