Package 'daMA'

June 26, 2025

Title Efficient design and analysis of factorial two-colour microarray data

Version 1.81.0

Date 1. October 2003

Author Jobst Landgrebe <jlandgr1@gwdg.de> and Frank Bretz <bretz@bioinf.uni-hannover.de>

Description This package contains functions for the efficient design of factorial two-colour microarray experiments and for the statistical analysis of factorial microarray data. Statistical details are described in Bretz et al. (2003, submitted)

Maintainer Jobst Landgrebe <jlandgr1@gwdg.de>

Imports MASS, stats

License GPL (>= 2)

URL http://www.microarrays.med.uni-goettingen.de

biocViews Microarray, TwoChannel, DifferentialExpression

git_url https://git.bioconductor.org/packages/daMA

git_branch devel

git_last_commit 9d56a44

git_last_commit_date 2025-04-15

Repository Bioconductor 3.22

Date/Publication 2025-06-26

Contents

analyseMA 2
cinfo 3
cinfoB.AB 4
cmat 4
cmatB.AB 5
core 5
data.3x2 6
designMA 7
designs.basic 8
designs.composite 9
id.3x2 9

Index 11

`analyseMA` *ANALYSIS OF FACTORIAL MICROARRAY EXPERIMENTS*

Description

analyseMA is used for the ananlysis of factorial two-colour microarray experiments based on the experimental design, a user-defined matrix containing the experimental question in contrast form and a vector to discern vectorial contrasts from contrasts given in matrix form.

Usage

```
analyseMA( data, design, id, cmat, cinfo, padj=c("none","bonferroni","fdr"), tol=1e-06 )
```

Arguments

`data`	a matrix of size $G \times N$ containing the normalized and/or standardized data to be analyzed, where G is the number of spots under investigation and N is the number of arrays used in the experiment. The matrix should contain one row for each spot. The matrix should contain as many columns as arrays involved in the experiment, such that each column contains the data for one single array. The matrix should not contain any ID variables, which are entered separately. Missing values should be entered as NA.
`design`	the design matrix of size $N \times (K + 2)$, where K is the number of experimental conditions. This is the design matrix X known from linear model theory and its elements are typically 0, 1, or -1. A 0 means that the associated parameter does not apply for the corresponding observation (i.e., row). The first two columns are reserved for the two dyes and are usually filled up with 1 and -1, respectively.
`id`	an ID vector of length G for the identification of the spots.
`cmat`	a matrix describing the p experimental questions (contrasts) to be analysed in the experiment. The matrix can be composed of vectorial contrasts (a single row of the matrix) and of contrasts in matrix form (several rows of the matrix), e.g. an $A \times B$ interaction effect in a 3×2 design. All contrasts have to be combined into one matrix (using rbind for instance).
`cinfo`	a vector of length p describing the grouping of the contrast matrix rows in vector or matrix form. E.g. if the design matrix contains three contrasts in vector form, cinfo = rep(1,3), if it contains two vectorial contratst and one as matrix with three rows, cinfo=c(1,1,3).
`padj`	a quoted string indicating the multiplicity adjustment that should be used. "none" - no multiplicity adjustment, "bonferroni" - Bonferroni single step adjustment, "fdr" - linear step-up procedure of Benjamini and Hochberg.
`tol`	A value indicating the tolerance for contrast estimability check

Details

The analysis is perfomed separately for each spot. For each spot, arrays with NA values are dropped. Then, for each experimental question (either contrast vector or contrast matrix) a check on the estimabilty of the resulting linear function is done. If the linear function of interest is estimable, t- or F-tests (whichever is appropriate) are computed and the associated unadjusted $P-$values are computed. Multiplicity adjustment is done over the number of spots only.

Value

a $G \times (4p + 3)$ matrix with the following row-wise components.

(i)	the first column contains the ID
(ii)	columns 2 though p+1 contain the estimates of the linear function (in case of vectorial contrasts) or the dregrees of freedom for the quadratic form in the numerator (in case of contrasts given in matrix form and that F-tests are used), depending on cinfo.
(iii)	columns p+2 through 2p+1 contain the test statistics (either t- or F-tests, depending on cinfo)
(iv)	columns 2p+2 through 3p+1 contain the raw P-values, associated to the t- and F-tests
(v)	column 3p+2 contains the mean square error
(vi)	column 3p+3 contains the residual degrees of freedom
(vii)	columns 3p+4 through 4p+3 contain the multiplicity adjusted P-values, associated to the raw P-values, as long as a multiplicty adjustment method has been selected

Author(s)

Jobst Landgrebe (jlandgr1@gwdg.de) and Frank Bretz (bretz@bioinf.uni-hannover.de)

References

Bretz, F and Landgrebe J and Brunner E (2003):"Design and analysis of two colour factorial microarray experiments", submitted. http://www.microarrays.med.uni-goettingen.de/

Examples

```
## Not run:     result <- analyseMA( data=data.3x2, design=designs.composite$BSBSBS, id=id.3x2,
  cmat=cmatB.AB, cinfo=c(1,3), padj=c("fdr"), tol=1e-06 ) # analyse a dataset with
         # 30012 spots and 18 arrays. The design
         # is 3x2 with 3 replicates, the
         # contrasts of interest are the main effect
         # B and the interaction effect AxB.

## End(Not run)
```

---

cinfo *Vector indexing the matrix cmat*

Description

This vector is used to describe the structure of the rows of the contrast matrix cmat. The number of entries in cinfo mirrors the number of experimental questions. "1" indicates a contrast in vectorial form, integers n > 1 indicate n contrasts given in matrix form.

References

Bretz, F and Landgrebe J and Brunner E (2003):"Design and analysis of two colour factorial microarray experiments", submitted.

Examples

```
data(cinfo)
```

`cinfoB.AB` *Vector indexing the matrix cmatB.AB*

Description

This vector is used to describe the structure of the rows of the contrast matrix cmatB.AB Its first element (1) indicates that the first esperimental question (main effect B) is described by a single contrast in vectorial form. The second element (2) indicates that the second experimental question (interaction between A and B) is given by a contrast in matrix form.

References

Bretz, F and Landgrebe J and Brunner E (2003):"Design and analysis of two colour factorial microarray experiments", submitted.

Examples

```
data(cinfoB.AB)
```

`cmat` *Contrast matrix describing the experimental questions*

Description

This matrix of numerical constants describes the experimental question, p say. Each experimental question is described by a single contrast vector (a single row in cmat) or by a contrast matrix (several rows in cmat). The ordering of the columns corresponds to that of the associated design matrix X. Thus, typically the first two elements in a row of cmat are reserved for for the two dyes. E.g. to compare the two dyes, we set (-1, 1, 0, ..., 0).

Usage

```
data(cmat)
```

References

Bretz, F and Landgrebe J and Brunner E (2003):"Design and analysis of two colour factorial microarray experiments", submitted.

Examples

```
data(cmat)
## maybe str(cmat) ; plot(cmat) ...
```

cmatB.AB — *Contrast matrix describing the experimental questions*

Description

This matrix of numerical constants describes the experimental question, say p. Each experimental question is described by a single contrast vector (a single row in cmat) or by a contrast matrix (several rows in cmat). The ordering of the columns corresponds to that of the associated design matrix X. Thus, typically the first two elements in a row of cmat are reserved for for the two dyes. E.g. the first line of the matrix cmatB.AB describes the main effect B.

Usage

```
data(cmatB.AB)
```

References

Bretz, F and Landgrebe J and Brunner E (2003):"Design and analysis of two colour factorial microarray experiments", submitted.

Examples

```
data(cmatB.AB)
```

core — *Internal function of analyseMA*

Description

This internal function of analyseMA computes the statistics and estimators that are organised and given out by the main function analyseMA.

Usage

```
core(vector, design, cmat, cinfo, tol)
```

Arguments

vector	a simple help variable for the apply call
design	the design matrix of size $N \times (K + 2)$, where K is the number of experimental conditions. This is the design matrix X known from linear model theory and its elements are typically 0, 1, or -1. A 0 means that the associated parameter does not apply for the corresponding observation (i.e., row). The first two columns are reserved for the two dyes and are usually filled up with 1 and -1, respectively.
cmat	a matrix describing the p experimental questions (contrasts) to be analysed in the experiment. The matrix can be composed of vectorial contrasts (a single row of the matrix) and of contrasts in matrix form (several rows of the matrix), e.g. an $A \times B$ interaction effect in a 3×2 design. All contrasts have to be combined into one matrix (using rbind for instance).

cinfo	a vector of length p describing the grouping of the contrast matrix rows in vector or matrix form. E.g. if the design matrix contains three contrasts in vector form, cinfo = rep(1,3), if it contains two vectorial contratst and one as matrix with three rows, cinfo=c(1,1,3).
tol	A value indicating the tolerance for contrast estimability check

Author(s)

Jobst Landgrebe (jlandgr1@gwdg.de) and Frank Bretz (bretz@bioinf.uni-hannover.de)

References

Bretz, F and Landgrebe J and Brunner E (2003):"Design and analysis of two colour factorial microarray experiments", submitted. http://www.microarrays.med.uni-goettingen.de/

---

data.3x2 *3x2 microarray data*

Description

These are data from a microarray experiment in which the expression profiles of three cell lines were analysed with and without drug treatment using cDNA microarrays spotted with 300012 human cDNAs. The data matrix consists of 30012 rows and 18 columns. Each row represents one spot, each column corresponds to one microarray.

Usage

```
data(data.3x2)
```

Format

The format is: matrix of size 30012 x 18.

Source

cDNA microarray lab of the University of Goettingen, Germany. http://www.microarrays.med.uni-goettingen.de

Examples

```
data(data.3x2)
## maybe str(data.3x2) ; plot(data.3x2) ...
```

designMA *DESIGN OF FACTORIAL MICROARRAY EXPERIMENTS*

Description

designMA computes efficient factorial microarray experimental designs for two-colour microarrays based on a list of user-defined design matrices, a matrix describing the experimental questions (contrasts), a vector to discern vectorial contrasts from contrasts given in matrix form and a design optimality criterion.

Usage

```
designMA(design.list, cmat, cinfo, type = c("d", "e", "t"), tol = 1e-06)
```

Arguments

design.list	a named list of design matrices. Each design matrix should have nrow = number of arrays and ncol= number of experimental conditions. With p columns, the first two columns describe the dye labeling (green and red), the remaining columns describe the experimental conditions.
cmat	a matrix describing the experimental questions (contrasts) to be analysed in the experiment. The matrix can be composed of vectorial contrasts (a single row of the matrix) and of contrasts in matrix form (several rows of the matrix), e.g. an $A \times B$ interaction effect in a 3×2 design. All contrasts have to be combined into one matrix (using rbind for instance).
cinfo	a vector describing the grouping of the contrast matrix rows in vector or matrix form. E.g. if the design matrix contains three contrasts in vector form, cinfo = rep(1,3), if it contains two vectorial contratst and one as matrix with three rows, cinfo=c(1,1,3).
type	a quoted letter indicating the optimality criterion that shoul be used. "d" - determinant, "e" - eigenvalue, "t" - trace.
tol	A value indicating the tolerance for contrast estimability check.

Details

The choice of the optimality criterion influences the design defined as best. We propose the trace criterion because of its straightforward interpretability. For a detailed description of optimality criteria cf. Pukelsheim, F. "Optimal Design of Experiments", New York 1993.

Value

a list with the following components

alleff	a matrix giving the absolute efficiency values (cols) for each contrast (rows). NA if contrast is not estimatable.
alleffrel	a matrix giving the relative efficiency values (cols) for each contrast (rows). The values are obtained by dividing the absolute values by the by the maximal efficiency value for a given contrast. NA if contrast is not estimatable.
alleffave	a vector giving the average efficiency for each design over all contrasts.
effdesign	the name of the design with the highest alleffave value.
df	a vector with the degrees of freedom of the F-statistics obtained by the designs.

Author(s)

Jobst Landgrebe (jlandgr1@gwdg.de) and Frank Bretz (bretz@bioinf.uni-hannover.de)

References

Bretz, F and Landgrebe J and Brunner E (2003):"Design and analysis of two colour factorial microarray experiments", submitted. http://www.microarrays.med.uni-goettingen.de/

Examples

```
  ## Not run: designs.basic # look at typical basic designs
  ## Not run: designs.composite  #look at comlpex composite designs
  ## Not run: t.eff.3x2.B.AB <-  designMA(designs.composite,
             cmatB.AB,cinfoB.AB,type="t")# compute design efficiencies for
          # a \eqn{3 \times 2} factorial experiment
          # using 18 microarrays and asking for
          # the main effect B and the interaction effect \eqn{A \times B}

## End(Not run)
  ## Not run: t.eff.3x2.all <-  designMA(designs.composite,
            cmat,cinfo,type="t")

## End(Not run)       #compute design efficiencies design for
          # a \eqn{3 \times 2} factorial
          # experiment using 18
                                                          # microarrays and asking for
                                                                   # the the simple B
                                                           # effects, the main effects
                                                            # A, B and the interaction
                                                            # effect \eqn{A  \times B}
```

`designs.basic` *Basic designs for two-colour factorial 3 x 2 microarray data*

Description

A list of matrices describing basic designs for two-colour factorial microarray data of size 3 x 2. Matrix rows represent microarrays, matrix columns represent parameters.

Usage

```
data(designs.basic)
```

Format

List of matrices of size 6 x 9.

Details

The designs are abbreviated as in the paper (cf. source and references): BS - swap over B, AL - A loop, XL - crossed loop, CL - circle loop, RS - star swap, TL - triangular loop, CR - common reference.

Source

cDNA microarray lab of the University of Goettingen, Germany. http://www.microarrays.med.uni-goettingen.de

---

`designs.composite` *Composite designs for two-colour factorial 3 x 2 microarray data*

Description

A list of matrices describing composite designs for two-colour factorial microarray data of size 3 x 2 using 18 microarrays each. The design matrices are made up of basic designs. Matrix rows represent microarrays, matrix columns represent parameters.

Usage

```
data(designs.composite)
```

Format

List of 10 matrices of size 18 x 9.

Details

The matrix names reflect the basic designs they are made up from. The first two digits of the names abbreviated the first basic design, the second two the second design etc. The basic design abbreviations are: BS - swap over B, AL - A loop, XL - crossed loop, CL - circle loop, RS - star swap, TL - triangular loop, CR - common reference. BSBSBS is a tripled basic BS design, CLCLTL is a double circle loop design combined with a triangular design and so on.

Source

cDNA microarray lab of the University of Goettingen, Germany. http://www.microarrays.med.uni-goettingen.de

Examples

```
data(designs.composite)
```

`id.3x2` *A vector of length 30012 containing numeric identifiers of the genes from the microarray dataset data.3x2.*

Description

Cf. data.3x2

Usage

```
data(id.3x2)
```

Format

The format is: num [1:30012] 12 24 108 120 204 216 300 312 396 408 ...

Examples

```
data(id.3x2)
```

Index

∗ **datasets**
    cinfo, 3
    cinfoB.AB, 4
    cmat, 4
    cmatB.AB, 5
    data.3x2, 6
    designs.basic, 8
    designs.composite, 9
    id.3x2, 9
∗ **design**
    analyseMA, 2
    designMA, 7
∗ **models**
    analyseMA, 2
    designMA, 7

analyseMA, 2

cinfo, 3
cinfoB.AB, 4
cmat, 4
cmatB.AB, 5
core, 5

data.3x2, 6
designMA, 7
designs.basic, 8
designs.composite, 9

id.3x2, 9