Package ‘RepeatedHighDim’

April 25, 2024

Type Package

Title Methods for High-Dimensional Repeated Measures Data

Version 2.3.0

Author Klaus Jung [aut, cre],
Jochen Kruppa [aut],
Sergej Ruff [aut]

Maintainer Klaus Jung <klaus.jung@tiho-hannover.de>

Description A toolkit for the analysis of high-dimensional repeated measurements, providing functions for outlier detection, differential expression analysis, gene-set tests, and binary random data generation.

License GPL (>= 3)

URL https://software.klausjung-lab.de

Encoding UTF-8

RoxygenNote 7.3.1

Imports ddaalpha, geometry, graphics, grDevices, MASS, mvtnorm, netmeta, nlme, progress, rgl, stats, utils

Suggests BiocManager, invgamma, limma

NeedsCompilation no

Repository CRAN

Date/Publication 2024-04-25 13:40:02 UTC

R topics documented:

- bag
- depmed
- fc_ci
- fc_plot
- GA_diagplot
- gem
- GlobTestMissing
Calculates the bag

Description
Calculates the bag of a gemplot (i.e. the inner gemstone).

Usage

bag(D, G)

Arguments
D
Data set with rows representing the individuals and columns representing the features. In the case of three dimensions, the colnames of D must be c("x", "y", "z").

G
List containing the grid information produced by gridfun and the halfspace location depths calculated by hldepth.

Details
Determines those grid points that belong to the bag, i.e. a convex hull that contains 50 percent of the data. In the case of a 3-dimensional data set, the bag can be visualized by an inner gemstone that can be accompanied by an outer gemstone (loop).
Value

A list containing the following elements:

- **coords** Coordinates of the grid points that belong to the bag. Each row represents a grid point and each column represents one dimension.

- **hull** A data matrix that contains the indices of the margin grid points of the bag that cover the convex hull by triangles. Each row represents one triangle. The indices correspond to the rows of coords.

Author(s)

Jochen Kruppa, Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

Examples

```r
## Attention: calculation is currently time-consuming.

## Not run:
## Two 3-dimensional example data sets D1 and D2
n <- 200
x1 <- rnorm(n, 0, 1)
y1 <- rnorm(n, 0, 1)
z1 <- rnorm(n, 0, 1)
D1 <- data.frame(cbind(x1, y1, z1))
x2 <- rnorm(n, 1, 1)
y2 <- rnorm(n, 1, 1)
z2 <- rnorm(n, 1, 1)
D2 <- data.frame(cbind(x2, y2, z2))
colnames(D1) <- c("x", "y", "z")
colnames(D2) <- c("x", "y", "z")

# Placing outliers in D1 and D2
D1[17, ] = c(4, 5, 6)
D2[99, ] = c(3, 4, 5)

# Grid size and graphic parameters
grid.size <- 20
```
depmed <- rgb(200, 100, 100, alpha = 100, maxColorValue = 255)
blue <- rgb(100, 100, 200, alpha = 100, maxColorValue = 255)
yel <- rgb(255, 255, 102, alpha = 100, maxColorValue = 255)
white <- rgb(255, 255, 255, alpha = 100, maxColorValue = 255)

require(rgl)
material3d(color=c(red, blue, yel, white),
alpha=c(0.5, 0.5, 0.5, 0.5), smooth=FALSE, specular=“black”)

# Calculation and visualization of gemplot for D1
G <- gridfun(D1, grid.size=20)
G$H <- hldepth(D1, G, verbose=TRUE)
dm <- depmed(G)
B <- bag(D1, G)
L <- loop(D1, B, dm=dm)
bg3d(color = “gray39”)
points3d(D1[L$outliers==0,1], D1[L$outliers==0,2], D1[L$outliers==0,3], col=“green”)
text3d(D1[L$outliers==1,1], D1[L$outliers==1,2], D1[L$outliers==1,3],
as.character(which(L$outliers==1)), col=yel)
spheres3d(dm[1], dm[2], dm[3], col=yel, radius=0.1)
material3d(1, alpha=0.4)
gem(B$coords, B$hull, red)
gem(L$coords.loop, L$hull.loop, red)
axes3d(col=“white”)

# Calculation and visualization of gemplot for D2
G <- gridfun(D2, grid.size=20)
G$H <- hldepth(D2, G, verbose=TRUE)
dm <- depmed(G)
B <- bag(D2, G)
L <- loop(D2, B, dm=dm)
points3d(D2[L$outliers==0,1], D2[L$outliers==0,2], D2[L$outliers==0,3], col=“green”)
text3d(D2[L$outliers==1,1], D2[L$outliers==1,2], D2[L$outliers==1,3],
as.character(which(L$outliers==1)), col=yel)
spheres3d(dm[1], dm[2], dm[3], col=yel, radius=0.1)
gem(B$coords, B$hull, blue)
gem(L$coords.loop, L$hull.loop, blue)

## End(Not run)

---

depmed

### Calculates the depth median.

#### Description

Calculates the depth median.

#### Usage

```
depmed(G)
```
Arguments

G List containing the grid information produced by `gridfun` and the halfspace location depths produced by `hldepth`.

Details

Calculates the depth median in a specified grid array with given halfspace location depth at each grid location.

Value

An vector with a length equal to the number of dimension of the array in G, containing the coordinates of the depth median.

Author(s)

Jochen Kruppa, Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: [https://software.klausjung-lab.de/](https://software.klausjung-lab.de/).

Examples

```r
## Attention: calculation is currently time-consuming.
## Not run:

# A 3-dimensional example data set D1
n <- 200
x1 <- rnorm(n, 0, 1)
y1 <- rnorm(n, 0, 1)
z1 <- rnorm(n, 0, 1)
D1 <- data.frame(cbind(x1, y1, z1))
colnames(D1) <- c("x", "y", "z")

# Specification of the grid and calculation of the halfspace location depth at each grid location.
G <- gridfun(D1, grid.size=20)
G$H <- hldepth(D1, G, verbose=TRUE)
dm <- depmed(G) ## Calculation of the depth median

## End(Not run)
```
**fc_ci**  
*Calculation of adjusted confidence intervals*

**Description**  
Calculation of adjusted confidence intervals

**Usage**  
```
fc_ci(fit, alpha = 0.05, method = "raw")
```

**Arguments**
- `fit` Object as returned from the function eBayes of the limma package
- `alpha` 1 - confidence level (e.g., if confidence level is 0.95, alpha is 0.05)
- `method` Either 'raw' for unadjusted confidence intervals, or 'BH' for Benjamini Hochberg-adjusted confidence intervals, or 'BY' for Benjamini Yekutieli-adjusted confidence intervals

**Details**  
Calculation of unadjusted and adjusted confidence intervals for the log fold change.

**Value**  
A results matrix with one row per gene, and one column for the p-value, the log fold change, the lower limit of the CI, and the upper limit of the CI.

**Author(s)**
Klaus Jung

**References**

**See Also**
For more information, please refer to the package’s documentation and the tutorial: [https://software.klausjung-lab.de/](https://software.klausjung-lab.de/).
Examples

```r
### Artificial microarray data
d = 1000  ### Number of genes
n = 10    ### Sample per group
fc = rlnorm(d, 0, 0.1)
mu1 = rlnorm(d, 0, 1)  ### Mean vector group 1
mu2 = mu1 * fc          ### Mean vector group 2
sd1 = rnorm(d, 1, 0.2)
sd2 = rnorm(d, 1, 0.2)
X1 = matrix(NA, d, n)  ### Expression levels group 1
X2 = matrix(NA, d, n)  ### Expression levels group 2
for (i in 1:n) {
  X1[,i] = rnorm(d, mu1, sd=sd1)
  X2[,i] = rnorm(d, mu2, sd=sd2)
}
X = cbind(X1, X2)
heatmap(X)

### Differential expression analysis with limma
if(check_limma()){
  group = gl(2, n)
design = model.matrix(~ group)
  fit1 = limma::lmFit(X, design)
  fit = limma::eBayes(fit1)
  ### Calculation of confidence intervals
  CI = fc_ci(fit=fit, alpha=0.05, method="raw")
  head(CI)
  CI = fc_ci(fit=fit, alpha=0.05, method="BH")
  head(CI)
  CI = fc_ci(fit=fit, alpha=0.05, method="BY")
  head(CI)
  fc_plot(CI, xlim=c(-0.5, 3), ylim=-log10(c(1, 0.0001)), updown="up")
  fc_plot(CI, xlim=c(-3, 0.5), ylim=-log10(c(1, 0.0001)), updown="down")
  fc_plot(CI, xlim=c(-3, 3), ylim=-log10(c(1, 0.0001)), updown="all")
}
```

---

fc_plot

**Volcano plot of adjusted confidence intervals**

Description

Volcano plot of adjusted confidence intervals

Usage

```r
fc_plot(CI,
```
alpha = 0.05,
updown = "all",
xlim = c(-3, 3),
ylim = -log10(c(1, 0.001))
)

Arguments

CI Object as returned from the function fc_ci
alpha 1 - confidence level (e.g., if confidence level is 0.95, alpha is 0.05)
updown Character, 'all' if CIs for all genes, 'down' if CIs for down-regulated genes, or 'up' if CIs for up-regulated genes to be plotted
xlim Vector of length 2 with the lower and upper limits for the X-axis
ylim Vector of length 2 with the lower and upper limits for the Y-axis. Please note, that p-values are usually displayed on the -log10-scale in a volcano plot

Details

Volcano plot of adjusted confidence intervals

Author(s)
Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

Examples

### Artificial microarray data
d = 1000 ### Number of genes
n = 10 ### Sample per group
fc = rlnorm(d, 0, 0.1)
mu1 = rlnorm(d, 0, 1) ### Mean vector group 1
mu2 = mu1 * fc ### Mean vector group 2
sd1 = rnorm(d, 1, 0.2)
scd2 = rnorm(d, 1, 0.2)
X1 = matrix(NA, d, n) ### Expression levels group 1
X2 = matrix(NA, d, n) ### Expression levels group 2
for (i in 1:n) {
    X1[,i] = rnorm(d, mu1, sd=sd1)
    X2[,i] = rnorm(d, mu2, sd=sd2)
}
X = cbind(X1, X2)
heatmap(X)

### Differential expression analysis with limma
if(check_limma()){  
group = gl(2, n)
design = model.matrix(~ group)
fit1 = limma::lmFit(X, design)
fit = limma::eBayes(fit1)

### Calculation of confidence intervals
CI = fc_ci(fit=fit, alpha=0.05, method="raw")
head(CI)
CI = fc_ci(fit=fit, alpha=0.05, method="BH")
head(CI)
CI = fc_ci(fit=fit, alpha=0.05, method="BY")
head(CI)

fc_plot(CI, xlim=c(-0.5, 3), ylim=-log10(c(1, 0.0001)), updown="up")
fc_plot(CI, xlim=c(-3, 0.5), ylim=-log10(c(1, 0.0001)), updown="down")
fc_plot(CI, xlim=c(-3, 3), ylim=-log10(c(1, 0.0001)), updown="all")
}

---

**GA_diagplot**

*Diagnostic plot for comparison of two correlation matrices.*

**Description**

A diagnostic plot that compares the entries of two correlation matrices using a color scale.

**Usage**

```r
GA_diagplot(
    R,
    Rt,
    eps = 0.05,
    col.method = "trafficlight",
    color = c(0, 8),
    top = ""
)
```
Arguments

R Specified correlation matrix.

Rt Correlation matrix of the data generated by the genetic algorithm.

eps Permitted difference between the entries of two matrices. Must only be specified if col.method="trafficlight".

col.method Method to use for color scaling the difference between the matrices. If method="trafficlight" only two colors are used, indicating whether the entries deviated at least by a difference of eps. If method="updown" a discrete gray scale is used.

color Value of two color that are used if method="trafficlight"

top Specifies the main title of the plot

Details

A diagnostic plot that compares the entries of two correlation matrices using a color scale.

Author(s)

Jochen Kruppa, Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

Examples

```r
## Not run:
R1 = diag(10)
X0 <- start_matrix(p=c(0.4, 0.2, 0.5, 0.15, 0.4, 0.35, 0.2, 0.25, 0.3, 0.4), k = 5000)
Xt <- iter_matrix(X0, R = diag(10), T = 10000, e.min = 0.00001)
GA_diagplot(R1, Rt = Xt$Rt, col.method = "trafficlight")
GA_diagplot(R1, Rt = Xt$Rt, col.method = "updown")
## End(Not run)
```
Plots a gemstone to an interactive graphics device.

Usage

gem(coords, hull, clr)

Arguments

- **coords**: Matrix with coordinates of the grid or of data points that belong to the gemstone, calculated by either `bag` or `loop`. Each row represents a grid point and each column represents one dimension.

- **hull**: Matrix with indices of triangles that cover a convex hull around the gemstone. Each row represents one triangle and the indices refer to the rows of `coords`.

- **clr**: Specifies the color of the gemstone.

Details

Only applicable to 3-dimensional data sets. Transparent colors are recommended for outer gemstone of the gemplot. Further graphical parameters can be set using `material3d()` of the rgl-package.

Author(s)

Jochen Kruppa, Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.
Examples

## Attention: calculation is currently time-consuming.
## Not run:

# Two 3-dimensional example data sets D1 and D2
n <- 200
x1 <- rnorm(n, 0, 1)
y1 <- rnorm(n, 0, 1)
z1 <- rnorm(n, 0, 1)
D1 <- data.frame(cbind(x1, y1, z1))
x2 <- rnorm(n, 1, 1)
y2 <- rnorm(n, 1, 1)
z2 <- rnorm(n, 1, 1)
D2 <- data.frame(cbind(x2, y2, z2))
colnames(D1) <- c("x", "y", "z")
colnames(D2) <- c("x", "y", "z")

# Placing outliers in D1 and D2
D1[17,] = c(4, 5, 6)
D2[99,] = -c(3, 4, 5)

# Grid size and graphic parameters
grid.size <- 20
red <- rgb(200, 100, 100, alpha = 100, maxColorValue = 255)
blue <- rgb(100, 100, 200, alpha = 100, maxColorValue = 255)
yel <- rgb(255, 255, 102, alpha = 100, maxColorValue = 255)
white <- rgb(255, 255, 255, alpha = 100, maxColorValue = 255)
require(rgl)
material3d(color=c(red, blue, yel, white),
alpha=c(0.5, 0.5, 0.5, 0.5), smooth=FALSE, specular="black")

# Calculation and visualization of gemplot for D1
G <- gridfun(D1, grid.size=20)
G$H <- hldepth(D1, G, verbose=TRUE)
dm <- depmed(G)
B <- bag(D1, G)
L <- loop(D1, B, dm=dm)
bg3d(color = "gray39")
points3d(D1[L$outliers==0,1], D1[L$outliers==0,2], D1[L$outliers==0,3], col="green")
text3d(D1[L$outliers==1,1], D1[L$outliers==1,2], D1[L$outliers==1,3],
as.character(which(L$outliers==1)), col=yel)
spheres3d(dm[1], dm[2], dm[3], col=yel, radius=0.1)
m3d(color = "gray39")
gem(B$coords, B$hull, red)
gem(L$coords.loop, L$hull.loop, red)
axes3d(col="white")

# Calculation and visualization of gemplot for D2
G <- gridfun(D2, grid.size=20)
G$H <- hldepth(D2, G, verbose=TRUE)
dm <- depmed(G)
B <- bag(D2, G)
```r
set.seed(123)
n <- 200
x1 <- rnorm(n, 0, 1)
x2 <- rnorm(n, 0, 1)
x3 <- rnorm(n, 0, 1)
x4 <- rnorm(n, 0, 1)
D <- data.frame(cbind(x1, x2, x3, x4))
D[67,] = c(7, 0, 0, 0)

G = gridfun(D, 20, 4)
G$H = hldepth(D, G, verbose=TRUE)
dm = depmed(G)
B = bag(D, G)
L = loop(D, B, dm=dm)
which(L$outliers==1)
date()
```

---

**GlobTestMissing**  
Detection of global group effect

**Description**  
Detection of global group effect

**Usage**  
`GlobTestMissing(X1, X2, nperm = 100)`

**Arguments**

- **X1**: Matrix of expression levels in first group. Rows represent features, columns represent samples.
- **X2**: Matrix of expression levels in second group. Rows represent features, columns represent samples.
- **nperm**: Number of permutations.
Details
Tests a global effect for a set of molecular features (e.g. genes, proteins,...) between the two groups of samples. Missing values are allowed in the expression data. Samples of the two groups are supposed to be unpaired.

Value
The p-value of a permutation test.

Author(s)
Klaus Jung

References

See Also
For more information, please refer to the package’s documentation and the tutorial: [https://software.klausjung-lab.de/](https://software.klausjung-lab.de/).

Examples
```r
### Global comparison of a set of 100 proteins between two experimental groups,
### where (tau * 100) percent of expression levels are missing.
n1 = 10
n2 = 10
d = 100
tau = 0.1
X1 = t(matrix(rnorm(n1*d, 0, 1), n1, d))
X2 = t(matrix(rnorm(n2*d, 0.1, 1), n2, d))
X1[sample(1:(n1*d), tau * (n1*d))] = NA
X2[sample(1:(n2*d), tau * (n2*d))] = NA
GlobTestMissing(X1, X2, nperm=100)
```

gridfun
Specifies grid for the calculation of the halfspace location depths

Description
Specifies a k-dimensional array as grid for the calculation of the halfspace location depths.

Usage
```r
gridfun(D, grid.size, k = 4)
```
hldepth

Arguments

D Data set with rows representing the individuals and columns representing the features. In the case of three dimensions, the colnames of D must be c("x", "y", "z").

grid.size Number of grid points in each dimension.

k Number of dimensions of the grid. Needs only be specified if D has more than columns.

Details

D must have at least three columns. If D has three columns, automatically a 3-dimensional grid is generated. If D has more than three columns, k must be specified.

Value

A list containing the following elements:

**H** The k-dimensional array.

In the case of a 3-dimensional array, additional elements are:

**grid.x, grid.y, grid.z** The coordinates of the grid points at each dimension.

In the case that the array has more than three dimensions, additional elements are:

**grid.k** A matrix with the coordinates of the grid. Row represents dimensions and columns represent grid points.

Author(s)

Jochen Kruppa, Klaus Jung

See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

---

**hldepth** Calculates the halfspace location depth

**Description**

Calculates the halfspace location depth for each point in a given grid.

**Usage**

hldepth(D, G, verbose = TRUE)
Arguments

D  Data set with rows representing the individuals and columns representing the features. In the case of three dimensions, the colnames of D must be c("x", "y", "z").

G  List containing the grid information produced by gridfun.

verbose  Logical. Indicates whether progress information is printed during calculation.

Details

Calculation of the halfspace location depth at each grid point is mandatory before calculating the depth median (depmed), the bag (bag) and the loop (loop). Ideally, the output is assigned to the array H produced by gridfun.

Value

H  An array of the same dimension as the array in argument G. The elements contain the halfspace location depth at the related grid location.

Author(s)

Jochen Kruppa, Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

Examples

## Attention: calculation is currently time-consuming.
## Not run:

# A 3-dimensional example data set D1
n <- 200
x1 <- rnorm(n, 0, 1)
y1 <- rnorm(n, 0, 1)
z1 <- rnorm(n, 0, 1)
D1 <- data.frame(cbind(x1, y1, z1))
colnames(D1) <- c("x", "y", "z")

# Specification of the grid and calculation of the halfspace location depth at each grid location.
G <- gridfun(D1, grid.size=20)
G$H <- hldepth(D1, G, verbose=TRUE)

## End(Not run)
Description

Starts the genetic algorithm based on a start matrix with specified marginal probabilities.

Usage

iter_matrix(X0, R, T = 1000, e.min = 1e-04, plt = TRUE, perc = TRUE)

Arguments

X0  
Start matrix with specified marginal probabilities. Can be generated by start_matrix.

R  
Desired correlation matrix the data should have after running the genetic algorithm.

T  
Maximum number of iterations after which the genetic algorithm stops.

e.min  
Minimum error (RMSE) between the correlation of the iterated data matrix and R.

plt  
Boolean parameter that indicates whether to plot e.min versus the iteration step.

perc  
Boolean parameter that indicates whether to print the percentage of iteration steps relativ to T.

Details

In each step, the genetic algorithm swaps two randomly selected entries in each column of X0. Thus it can be guaranteed that the marginal probabilities do not change. If the correlation matrix is closer to R than that of X0(t-1), X0(t) replaces X0(t-1).

Value

A list with four entries:

Xt Final representativ data matrix with specified marginal probabilities and a correlation as close as possible to R

t Number of performed iteration steps (t <= T)

Rt Empirical correlation matrix of Xt

RMSE Final RSME error between desired and achieved correlation matrix

Author(s)

Jochen Kruppa, Klaus Jung
References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klauskjung-lab.de/.

Examples

```r
### Generation of the representive matrix Xt
X0 <- start_matrix(p = c(0.5, 0.6), k = 1000)
Xt <- iter_matrix(X0, R = diag(2), T = 10000, e.min = 0.00001)$Xt

### Drawing of a random sample S of size n = 10
S <- Xt[sample(1:1000, 10, replace = TRUE),]
```

---

**loop**

*Calculates the fence and the loop*

**Description**

Calculates the fence and the loop of a gemplot (i.e. the outer gemstone).

**Usage**

```r
loop(D, B, inflation = 3, dm)
```

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>Data set with rows representing the individuals and columns representing the features. In the case of three dimensions, the colnames of D must be c(&quot;x&quot;, &quot;y&quot;, &quot;z&quot;).</td>
</tr>
<tr>
<td>B</td>
<td>List containing the information about the coordinates of the bag and the convex hull that forms the bag (determined by <code>bag</code>).</td>
</tr>
<tr>
<td>inflation</td>
<td>A numeric value &gt; 0 that specifies the inflation factor of the bag relative to the median (default = 3).</td>
</tr>
<tr>
<td>dm</td>
<td>The coordinates of the depth median as produced by <code>depmed</code>.</td>
</tr>
</tbody>
</table>

**Details**

The fence inflates the the bag relative to the depth median by the factor `inflation`. Data points outside the bag and inside the fence the loop or outer gemstone are flagged as outliers. Data points outside the fence are marked as outliers. In the case of a 3-dimensional data set, the loop can be visualized by an outer gemstone around the inner gemstone or bag.
### Value

A list containing the following elements:

- **coords.loop** Coordinates of the data points that are inside the convex hull around the loop.
- **hull.loop** A data matrix that contains the indices of the margin data points of the loop that cover the convex hull by triangles. Each row represents one triangle. The indices correspond to the rows of coords.loop.
- **coords.fence** Coordinates of the grid points that are inside the fence but outside the bag.
- **hull.fence** A data matrix that contains the indices of the margin grid points of the fence that cover the convex hull around the fence by triangles. Each row represents one triangle. The indices correspond to the rows of coords.fence.
- **outliers** A vector of length equal to the sample size. Data points that are inside the fence are labelled by 0 and values outside the fence (i.e. outliers) are labelled by 1.

### Author(s)

Jochen Kruppa, Klaus Jung

### References


### See Also

For more information, please refer to the package’s documentation and the tutorial: [https://software.klausjung-lab.de/](https://software.klausjung-lab.de/).

### Examples

```r
## Attention: calculation is currently time-consuming.
## Not run:

# Two 3-dimensional example data sets D1 and D2
n <- 200
x1 <- rnorm(n, 0, 1)
y1 <- rnorm(n, 0, 1)
z1 <- rnorm(n, 0, 1)
D1 <- data.frame(cbind(x1, y1, z1))
x2 <- rnorm(n, 1, 1)
y2 <- rnorm(n, 1, 1)
z2 <- rnorm(n, 1, 1)
D2 <- data.frame(cbind(x2, y2, z2))
colnames(D1) <- c("x", "y", "z")
colnames(D2) <- c("x", "y", "z")
```

# Placing outliers in D1 and D2
D1[,1] = c(4, 5, 6)
D2[,99] = c(3, 4, 5)

# Grid size and graphic parameters
grid.size <- 20
red <- rgb(200, 100, 100, alpha = 100, maxColorValue = 255)
blue <- rgb(100, 100, 200, alpha = 100, maxColorValue = 255)
yel <- rgb(255, 255, 102, alpha = 100, maxColorValue = 255)
white <- rgb(255, 255, 255, alpha = 100, maxColorValue = 255)
require(rgl)
material3d(color=c(red, blue, yel, white),
 alpha=c(0.5, 0.5, 0.5, 0.5), smooth=FALSE, specular="black")

# Calculation and visualization of gemplot for D1
G <- gridfun(D1, grid.size=20)
G$H <- hldepth(D1, G, verbose=TRUE)
dm <- depmed(G)
B <- bag(D1, G)
L <- loop(D1, B, dm=dm)
bg3d(color = "gray39")
points3d(D1[L$outliers==0,1], D1[L$outliers==0,2], D1[L$outliers==0,3], col="green")
text3d(D1[L$outliers==1,1], D1[L$outliers==1,2], D1[L$outliers==1,3],
 as.character(which(L$outliers==1)), col=yel)
spheres3d(dm[1], dm[2], dm[3], col=yel, radius=0.1)
material3d(1, alpha=0.4)
gem(B$coords, B$hull, red)
gem(L$coords.loop, L$hull.loop, red)
axes3d(col="white")

# Calculation and visualization of gemplot for D2
G <- gridfun(D2, grid.size=20)
G$H <- hldepth(D2, G, verbose=TRUE)
dm <- depmed(G)
B <- bag(D2, G)
L <- loop(D2, B, dm=dm)
points3d(D2[L$outliers==0,1], D2[L$outliers==0,2], D2[L$outliers==0,3], col="green")
text3d(D2[L$outliers==1,1], D2[L$outliers==1,2], D2[L$outliers==1,3],
 as.character(which(L$outliers==1)), col=yel)
spheres3d(dm[1], dm[2], dm[3], col=yel, radius=0.1)
gem(B$coords, B$hull, blue)
gem(L$coords.loop, L$hull.loop, blue)

## End(Not run)
Description

This function conducts network meta-analysis using gene expression data to make indirect comparisons between different groups. It computes the p values for each gene and the fold changes, and provides a dataframe containing these results.

Usage

netRNA(TE, seTE, treat1, treat2, studlab)

Arguments

**TE**
A list containing log fold changes from two individual studies. Index names of the list should be the gene names; otherwise, each value of the 'name' column in the output dataframe will correspond to the position in the list, rather than gene identifiers.

**seTE**
A list containing standard errors of log fold changes from two individual studies.

**treat1**
A vector with Label/Number for first treatment.

**treat2**
A vector with Label/Number for second treatment.

**studlab**
A vector containing study labels

Details

The function supports a simple network with three nodes, where one node represents a control group and the two other nodes represent treatment (or diseased) groups. While the user provides fold changes and their standard errors of each treatment versus control as input, the function calculates the fold changes for the indirect comparison between the two treatments. It’s crucial to note that the order of genes in the TE and seTE lists for both studies should be the same. Meaning if Gene "A" is the first gene in the first study, it should also be the first gene in the second study.

Value

A list containing the p values for each gene, the fold changes, the upper and lower bounds for the 95% CI of the log fold changes, and a summary dataframe with results for each gene.

Author(s)

Klaus Jung, Sergej Ruff

References


See Also

For more information, please refer to the package's documentation and the tutorial: https://software.klausjung-lab.de/.

Examples

# Data generation

### Basic expression, fold change, batch effects and error
alpha.1 = rnorm(G, 0, 1)
alpha.2 = rnorm(G, 0, 1)
beta.1 = rnorm(G, 0, 1)
beta.2 = rnorm(G, 0, 1)
gamma.1 = rnorm(G, 0, 1)
gamma.2 = rnorm(G, 2, 1)
delta.1 = sqrt(invgamma::rinvgamma(G, 1, 1))
delta.2 = sqrt(invgamma::rinvgamma(G, 1, 2))
sigma.g = rep(1, G)

# Generate gene names
gene_names <- paste("Gene", 1:G, sep = "")

### Data matrices of control and treatment (disease) groups
C.1 = matrix(NA, G, n)
C.2 = matrix(NA, G, n)
T.1 = matrix(NA, G, n)
T.2 = matrix(NA, G, n)

for (j in 1:n) {
  C.1[,j] = alpha.1 + (0 * beta.1) + gamma.1 + (delta.1 * rnorm(1, 0, sigma.g))
  C.2[,j] = alpha.1 + (0 * beta.2) + gamma.2 + (delta.2 * rnorm(1, 0, sigma.g))
  T.1[,j] = alpha.2 + (1 * beta.1) + gamma.1 + (delta.1 * rnorm(1, 0, sigma.g))
  T.2[,j] = alpha.2 + (1 * beta.2) + gamma.2 + (delta.2 * rnorm(1, 0, sigma.g))
}

study1 = cbind(C.1, T.1)
study2 = cbind(C.2, T.2)

# Assign gene names to row names
rownames(study1) <- gene_names
rownames(study2) <- gene_names

Differential Analysis

if(check_limma()){
  ### study1: treatment A versus control
  ...
RepeatedHighDim

RepeatedHighDim Package

Description

A comprehensive toolkit for repeated high-dimensional analysis.

Details

The RepeatedHighDim-package is a collection of functions for the analysis of high-dimensional repeated measures data, e.g. from Omics experiments. It provides function for outlier detection,
differential expression analysis, self-contained gene-set testing, and generation of correlated binary data.

For more information and examples, please refer to the package documentation and the tutorial available at https://software.klausjung-lab.de/.

## Functions

This package includes the following functions:

**B**
- **bag**: Calculates the bag.

**D**
- **depmed**: Calculates the depth median.

**F**
- **fc_ci**: Calculates adjusted confidence intervals.
- **fc_plot**: Creates a volcano plot of adjusted confidence intervals.

**G**
- **GA_diagplot**: Generates a diagnostic plot for comparing two correlation matrices.
- **gem**: Plots a gemstone to an interactive graphics device.
- **GlobTestMissing**: Detects global group effects.
- **gridfun**: Specifies a grid for calculating halfspace location depths.

**H**
- **hldepth**: Calculates the halfspace location depth.

**I**
- **iter_matrix**: Implements a genetic algorithm for generating correlated binary data.

**L**
- **loop**: Calculates the fence and the loop.

**N**
- **netRNA**: network meta-analysis using gene expression data.

**R**
- **RHighDim**: Detects global group effects.
- **rho_bounds**: Calculates lower and upper bounds for pairwise correlations.
- **rmvbinary_QA**: Simulates correlated binary variables using the algorithm by Qaqish (2003).
RHighDim

S:
- `sequence_probs`: Calculates probabilities for binary sequences.
- `start_matrix`: Sets up the start matrix.
- `summary_RHD`: Provides a summary of the RHighDim function.

T:
- `TestStatSimple`: Calculates the test statistic for RHighDim.
- `TestStatSP`: Calculates the test statistic for RHighDim.

Author(s)

Maintainer: Klaus Jung (<klaus.jung@tiho-hannover.de>)

Other contributors:

- Jochen Kruppa (<j.kruppa@hs-osnabrueck.de>)
- Sergej Ruff (<Sergej.Ruff@tiho-hannover.de>)

If you have any questions, suggestions, or issues, please feel free to contact the maintainer, Klaus Jung (<klaus.jung@tiho-hannover.de>).

See Also

For more information, please refer to the package’s documentation and the tutorial: [https://software.klausjung-lab.de/](https://software.klausjung-lab.de/).

---

RHighDim

**Detection of global group effect**

**Description**

Detection of global group effect

**Usage**

`RHighDim(X1, X2, paired = TRUE)`

**Arguments**

- `X1`: Matrix of expression levels in first group. Rows represent features, columns represent samples.
- `X2`: Matrix of expression levels in second group. Rows represent features, columns represent samples.
- `paired`: FALSE if samples are unpaired, TRUE if samples are paired.
Details

Global test for a set of molecular features (e.g. genes, proteins,...) between two experimental groups. Paired or unpaired design is allowed.

Value

An object that contains the test results. Contents can be displayed by the summary function.

Author(s)

Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

Examples

```r
### Global comparison of a set of 100 genes between two experimental groups.
X1 = matrix(rnorm(1000, 0, 1), 10, 100)
X2 = matrix(rnorm(1000, 0.1, 1), 10, 100)
RHD = RHighDim(X1, X2, paired=FALSE)
summary_RHD(RHD)
```

rho_bounds

Calculate lower and upper the bounds for pairwise correlations

**Description**

Calculate lower and upper the bounds for pairwise correlations

**Usage**

```
rho_bounds(R, p)
```

**Arguments**

- **R**  
  Correlation matrix
- **p**  
  Vector of marginal frequencies
Details
The function calculates upper and lower bounds for pairwise correlations given a vector of marginal probabilities as detailed in Emrich and Piedmonte (1991).

Value
A list with three entries:

- $L$: Matrix of lower bounds
- $U$: Matrix of upper bounds
- $Z$: Matrix that indicates whether specified correlations in $R$ are bigger or smaller than the calculated bounds

Author(s)
Jochen Kruppa, Klaus Jung

References

See Also
For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

Examples
```r
### A simple example
R <- diag(4)
p <- c(0.1, 0.2, 0.4, 0.5)
 rho_bounds(R, p)
```

Description
Generation of random sample of binary correlated variables

Usage
```r
rmvbinary_EP(n, R, p)
```
Arguments

\(n\)  Sample size
\(R\)  Correlation matrix
\(p\)  Vector of marginal probabilities

Details

The function implements the algorithm proposed by Emrich and Piedmonte (1991) to generate a random sample of \(d (=\text{length}(p))\) correlated binary variables. The sample is generated based on given marginal probabilities \(p\) of the \(d\) variables and their correlation matrix \(R\). The algorithm generates first determines an appropriate correlation matrix \(R'\) for the multivariate normal distribution. Next, a sample is drawn from \(N_d(0, R')\) and each variable is finally dichotomized with respect to \(p\).

Value

Sample \((n \times p)\)-matrix with representing a random sample of size \(n\) from the specified multivariate binary distribution.

Author(s)

Jochen Kruppa, Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

Examples

```r
## Generation of a random sample
rmvbinary_EP(n = 10, R = diag(2), p = c(0.5, 0.6))
```

Description

Generation of random sample of binary correlated variables
Usage

`rmvbinary_QA(n, R, p)`

Arguments

- `n` Sample size
- `R` Correlation matrix
- `p` Vector of marginal probabilities

Details

The function implements the algorithm proposed by Qaqish (2003) to generate a random sample of `d` (=length(p)) correlated binary variables. The sample is generated based on given marginal probabilities `p` of the `d` variables and their correlation matrix `R`. The algorithm starts by generating a data for the first variable `X_1` and generates successively the data for `X_2`, ... based on their conditional probabilities `P(X_j|X_{i-1},...,X_1)`, `j=1,...,d`.

Value

Sample `(n x p)`-matrix representing a random sample of size `n` from the specified multivariate binary distribution.

Author(s)

Jochen Kruppa, Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.

Examples

```r
## Generation of a random sample
rmvbinary_QA(n = 10, R = diag(2), p = c(0.5, 0.6))
```
sequence_probs

Calculation of probabilities for binary sequences

Description
Calculation of probabilities for binary sequences based on the final matrix generated by the genetic algorithm.

Usage
sequence_probs(Xt)

Arguments
Xt
Representative matrix generated by the genetic algorithm with iter_matrix.

Details
Observation of binary correlated binary data can be expressed as binary sequences. In the case of two binary variables possible observations are (0,0), (0,1), (1,0) and (1,1). In general, $2^m$ binary sequences are possible, where $m$ is the number of binary variables. Based on the representative matrix generated by the genetic algorithm the probability for each binary sequence is determined.

Value
A vector of probabilities for the binary sequences.

Author(s)
Jochen Kruppa, Klaus Jung

References

See Also
For more information, please refer to the package’s documentation and the tutorial: https://software.klausjung-lab.de/.
### start_matrix

#### Description

Generation of the start matrix with n rows and specified marginal probabilities p.

#### Usage

```r
start_matrix(p, k)
```

#### Arguments

- **p**
  - Marginal probabilities of the start matrix.
- **k**
  - Number of rows to be generated.

#### Details

The start matrix needs to be setup for further use in the genetic algorithm implemented in the function `iter_matrix`. For high-dimensional cases or if the marginal probabilities have multiple decimal places, the number k of rows should be large (up to multiple thousand).

#### Value

A (k x p)-Matrix with with entries 0 and 1 according to the specified marginal probabilities p.

#### Author(s)

Jochen Kruppa, Klaus Jung

#### References


#### See Also

For more information, please refer to the package’s documentation and the tutorial: [https://software.klausjung-lab.de/](https://software.klausjung-lab.de/).
Examples

```r
X0 <- start_matrix(p = c(0.5, 0.6), k = 10000)

### check if p can be restored
apply(X0, 2, mean)
```

summary_RHD

---

Summary of RHighDim function

Description

Summary of RHighDim function

Usage

```r
summary_RHD(object, ...)
```

Arguments

- `object` An object provided by the RHighDim function.
- `...` additional arguments affecting the summary produced.

Details

Summarizes the test results obtained by the RHighDim function.

Value

No value

Author(s)

Klaus Jung

References


See Also

For more information, please refer to the package’s documentation and the tutorial: [https://software.klausjung-lab.de/](https://software.klausjung-lab.de/).
Examples

### Global comparison of a set of 100 genes between two experimental groups.

```r
X1 = matrix(rnorm(1000, 0, 1), 10, 100)
X2 = matrix(rnorm(1000, 0.1, 1), 10, 100)
RHD = RHighDim(X1, X2, paired=FALSE)
summary_RHD(RHD)
```

---

TestStatSimple

Calculates test statistic for RHighDim

### Description

Calculates test statistic for RHighDim

#### Usage

```r
TestStatSimple(Y, H)
```

#### Arguments

- **Y**
  - A matrix.
- **H**
  - A matrix.

#### Value

A list.

#### See Also

For more information, please refer to the package’s documentation and the tutorial: [https://software.klausjung-lab.de/](https://software.klausjung-lab.de/).

---

TestStatSP

Calculates test statistic for RHighDim

### Description

Calculates test statistic for RHighDim

#### Usage

```r
TestStatSP(Y1, Y2)
```

#### Arguments

- **Y1**
  - A matrix.
- **Y2**
  - A matrix.
Value

A list.

See Also

For more information, please refer to the package's documentation and the tutorial: https://software.klausjung-lab.de/.
Index

bag, 2, 11, 16, 18, 24

depmed, 4, 16, 18, 24

fc.ci, 6, 24
fc.plot, 7, 24

GA.diagplot, 9, 24
gem, 11, 24
GlobTestMissing, 13, 24
gridfun, 2, 5, 14, 16, 24

hldepth, 2, 5, 15, 24

iter_matrix, 17, 24, 30, 31

loop, 2, 11, 16, 18, 24

netRNA, 20, 24

RepeatedHighDim, 23
RepeatedHighDim-package
(RepeatedHighDim), 23

RHighDim, 24, 25
rho_bounds, 24, 26
rmvbinary_EP, 24, 27
rmvbinary_QA, 24, 28

sequence_probs, 25, 30
start_matrix, 17, 25, 31
summary_RHD, 25, 32

TestStatSimple, 25, 33
TestStatSP, 25, 33