

Package ‘knorm’

February 9, 2012

Type Package

Title Microarray Data From Multiple Biologically Interrelated Experiments

Version 1.0

Date 2008-04-07

Author Siew-Leng Teng

Maintainer Siew Leng Teng <slteng@stat.berkeley.edu>

Depends R (>= 2.4.0), MASS

Description This package provides Knorm correlations between genes (or probes) from microarray data obtained across multiple biologically interrelated experiments. The Knorm correlation adjusts for experiment dependencies (correlations) and reduces to the Pearson coefficient when experiment dependencies are absent. The Knorm estimation approach can be generally applicable to obtain between-row correlations from data matrices with two-way dependencies.

License GPL (>= 2)

Repository CRAN

Date/Publication 2008-04-08 06:43:47

R topics documented:

knorm	2
mmcad	3

Index	5
--------------	----------

Description

Produces Knorm correlations between genes (or probes) from microarray data obtained across multiple biologically interrelated experiments. The Knorm correlation adjusts for experiment dependencies (correlations) and reduces to the Pearson coefficient when experiment dependencies are absent. The Knorm estimation approach can be generally applicable to obtain between-row correlations from data matrices with two-way dependencies.

Usage

```
knorm(data, bsamples, thres_diff, thres_ev1, thres_ev2, burn_in, no_subgenes, no_fullgenes, repli)
```

Arguments

data	matrix containing (normalized) gene expression data. Rows correspond to arrays and columns correspond to genes (or probes). Data from replicates of experiments are placed in consecutive rows.
bsamples	number of bootstrap samples for estimation.
thres_diff	threshold of difference between log-likelihood values.
thres_ev1	threshold for eigen values of experiment covariance matrix. Eigen values smaller than thres_ev1 are considered negligible.
thres_ev2	threshold for eigen values of gene covariance matrix. Eigen values smaller than thres_ev2 are considered negligible.
burn_in	minimum number of iterations for estimation. Default is 2.
no_subgenes	number of genes (or probes) to be used in the row-subsampling technique for estimating the experiment covariance matrix. This number should not be more than the number of experiments in the data.
no_fullgenes	number of genes (or probes) in data.
repli	vector of number of replicates for each experiment. For example, c(2,3) denotes two replicates for experiment 1 and three replicates for experiment 2.

Details

This estimation procedure consists of a gene (or row) sub-sampling and a covariance shrinkage technique that iteratively estimates the gene and experiment covariance matrices. The covariance shrinkage method using the diagonal matrix with unequal covariances as the target matrix was used (Schafer and Strimmer, 2005). For more details on the estimation procedure, model assumptions and conditions, please refer to Teng et al. (2007).

Value

A list containing:

a_cor_est	Experiment correlation matrix estimate.
g_cor_est	Knorm correlations (between genes).
m_est	Mean matrix estimate.

Author(s)

Siew Leng Teng

References

Teng, S.L., Huang, H., and Zhou, X. Jasmine. (2008), "A statistical framework to infer functional gene relationships from biologically interrelated microarray experiments"

Examples

```
#Importing simulated Multiple Microarray data.
#For more information on data set imported, look at help file for mmcd
#for futher information.
data(mmcd)

#Creating vector fo the number of replicates for each experiment. There
#will be three replications for each experiment in the mmcd data.
repli=rep(3,30)

results <- knorm(mmcd, 25, 0.01, 1e-10, 1e-10, 2, length(repli),ncol(mmcd),repli)
a_cor_est <- results$a_cor_est
g_cor_est <- results$g_cor_est
```

mmcad

Multiple Microarray Example Dataset - A simulated dataset example

Description

This is a simple simulated dataset of gene expression values for 100 genes (or probes) and 30 arrays, with 3 replicates for each array. Values are generated from a multivariate matrix normal distribution with a specified mean matrix, covariance matrices and other (nuisance) parameters (see Teng et al., 2007 for more details).

Usage

```
mmcad
```

Format

Data is represented in a matrix where rows represent arrays and columns represent genes. Values from replicated arrays are placed in consecutive rows.

References

Teng, S.L., Huang, H., and Zhou, X. Jasmine. (2008), "A statistical framework to infer functional gene relationships from biologically interrelated microarray experiments".

Examples

```
#Importing simulated Multiple Microarray data.
#For more information on data set imported, look at help file for mmcd
#for futher information.
data(mmcad)

#Creating vector fo the number of replicates for each experiment. There
#will be three replications for each experiment in the mmcd data.
repli=rep(3,30)

results <- knorm(mmcad, 25, 0.01, 1e-10, 1e-10, 2, length(repli),ncol(mmcad),repli)
a_cor_est <- results$a_cor_est
g_cor_est <- results$g_cor_est
```

Index

*Topic **datasets**

mmcad, [3](#)

*Topic **models**

knorm, [2](#)

knorm, [2](#)

mmcad, [3](#)