Package 'JGNI'

July 26, 2017

Type Package

Title Joint Gene Network Inference (JGNI)

Version 1.0

Date 2017-07-26

Author Xiao-Fei Zhang

Maintainer Xiao-Fei Zhang <zhangxf@mail.ccnu.edu.cn>

Description JGNI is a graphical model for joint gene network inference across multiple subpopulations and data types. JGNI can model both Gaussian and nonparanormal data.

Depends igraph, MASS, Matrix

License GPL-2

R topics documented:

`JGNI-package` *Joint Gene Network Inference*

Description

Run the graphical model for joint gene network inference across multiple subpopulations and data types.

Details

Package:	JGNI
Type:	Package
Version:	1.0
Date:	2017-07-26
License:	GPL (>= 2)
LazyLoad:	yes

Author(s)

Xiao-Fei Zhang
Maintainer Xiao-Fei Zhang <zhangxf@mail.ccnu.edu.cn>

References

Xiao-Fei Zhang, Le Ou-Yang, Shuo Yang, Hong Qin, Xiaohua Hu and Hong Yan (2017), JGNI: A graphical model for joint gene network inference across multiple subpopulations and data types.

Examples

```
# Simulation data
dat = generate_data(K = 3, G = 4, n = 100, p = 100, tau = 0.5, network.type = "ER")
result = JGNI(dat$X, 0.6, 0.3, model = "Gaussian", weights = "equal")

# TCGA breast cancer data
data("TCGA.BRCA")
result = JGNI(TCGA.BRCA$X, 1.1, 0.4, model = "nonparanormal", weights = "equal")
```

---

admm *ADMM algorithm for JGNI*

Description

ADMM algorithm for JGNI (Algorithm 1 in the Supplementary Information). For details, refer to Supplementary Section S3.1.

Usage

```
admm(S, lambda, alpha, n = NULL, penalize.diagonal = FALSE, epsilon = 1e-05, maxiter = 500, rho
```

Arguments

S	A matrix ($K \times G$) of list of sample covariance matrices ($p \times p$), where K is the number of data types and G is the number of subpopulations. The (k,g)-th element is a $p \times p$ sample covariance matrix corresponding to the k-th data type and the g-th subpopulation.
lambda	The tuning parameter for controlling the level of sparsity of networks.

alpha	The tuning parameter for controlling the extend of similarity among subpopulation-specific networks.
n	The sample size. A vector with length equal to the number of subpopulations.
penalize.diagonal	Determines whether the sparsity penalty is applied to the diagonal.
epsilon	The tolerance parameter for convergence criteria.
maxiter	The maximum number of iterations for the ADMM algorithm.
rho	The penalty parameter in the ADMM algorithm.
rho.incr	The increase step parameter for varying penalty parameter rho.
rho.max	The maximum value of rho.

Details

The function is used to solve the optimization model of JGNI.

Value

Omega.hat	A matrix ($K \times G$) of list of estimated precision matrices.
R.hat	A matrix ($K \times 1$) of list of estimated common components.
M.hat	A matrix ($K \times G$) of list of estimated subpopulation-unique components.

Author(s)

Xiao-Fei Zhang

References

Xiao-Fei Zhang, Le Ou-Yang, Shuo Yang, Hong Qin, Xiaohua Hu and Hong Yan (2017), JGNI: A graphical model for joint gene network inference across multiple subpopulations and data types.

See Also

JGNI, generate_data, Corr, TCGA.BRCA

Examples

```
# Simulation data
dat = generate_data(K = 3, G = 4, n = 100, p = 100, tau = 0.5, network.type = "ER")
S = matrix(list(), 3, 4)
for (k in 1:3){
  for(g in 1:4){
    S[[k,g]] = Corr(dat$X[[k,g]], method = "pearson")
  }
}
result = admm(S, 0.6, 0.3)

# TCGA breast cancer data
data("TCGA.BRCA")
result = admm(TCGA.BRCA$Sigma, 1.1, 0.4)
```

Corr *Compute the sample covariance (or correlation) matrix*

Description

Compute the sample covariance (or correlation) matrix for a data matrix.

Usage

```
Corr(X, method = "pearson")
```

Arguments

X	A $n \times p$ data matrix.
method	A character string indicating which correlation coefficient (or covariance) is to be computed. "pearson" (default) and "kendall" can be used.

Details

The function is used to compute sample covariance (or correlation) matrix for a given data matrix.

Value

S	A $p \times p$ sample covariance (or correlation) matrix.

Author(s)

Xiao-Fei Zhang

References

Xiao-Fei Zhang, Le Ou-Yang, Shuo Yang, Hong Qin, Xiaohua Hu and Hong Yan (2017), JGNI: A graphical model for joint gene network inference across multiple subpopulations and data types.

See Also

JGNI, admm, generate_data, TCGA.BRCA

Examples

```
# Simulation data
dat = generate_data(K = 3, G = 4, n = 100, p = 100, tau = 0.5, network.type = "ER")
S = matrix(list(), 3, 4)
for (k in 1:3){
  for(g in 1:4){
    S[[k,g]] = Corr(dat$X[[k,g]], method = "pearson")
  }
}

# TCGA breast cancer data
data("TCGA.BRCA")
S =  Corr(TCGA.BRCA$X[[1,1]], method = "kendall")
```

generate_data *Generate simulation data*

Description

The complete procedure for generating simulation data (Algorithm 3 in the Supplementary Information). For details, refer to Supplementary Section S3.6.

Usage

```
generate_data(K = 3, G = 4, n = 100, p = 100, tau = 0.5, network.type = "ER", umin = 0.5, umax =
```

Arguments

K	The number of data types.
G	The number of subpopulations.
n	The sample size. A positive intege or a vector with length equal to the number of subpopulations.
p	The number of genes.
tau	The ratio of the number of subpopulation-unique edges to the number of common edges.
network.type	A character string indicating which network type is generated. "ER" (Erdos-Renyi) and "SF" (scale-free) can be used.
umin	The lower limits of the edge values.
umax	The upper limits of the edge values.

Details

The function is used to generate simulation data.

Value

X	A matrix ($K \times G$) of list of data matrices ($n_g \times p$), where K is the number of data types and G is the number of subpopulations. The (k,g)-th element is a $n_g \times p$ data matrix corresponding to the k-th data type and the g-th subpopulation.
Omega	A matrix ($K \times G$) of list of the precision matrices.
R	A matrix ($K \times 1$) of list of the common components.
M	A matrix ($K \times G$) of list of the subpopulation-unique components.

Author(s)

Xiao-Fei Zhang

References

Xiao-Fei Zhang, Le Ou-Yang, Shuo Yang, Hong Qin, Xiaohua Hu and Hong Yan (2017), JGNI: A graphical model for joint gene network inference across multiple subpopulations and data types.

See Also

JGNI, admm, Corr, TCGA.BRCA

Examples

```
# Simulation data
dat = generate_data(K = 3, G = 4, n = 100, p = 100, tau = 0.5, network.type = "ER")
```

JGNI *Joint gene network inference across multiple subpopulations and data types*

Description

The complete procedure for inferring gene networks from gene expression data using JGNI (Algorithm 1 in the Supplementary Information). For details, refer to Supplementary Section S3.2.

Usage

```
JGNI(X, lambda, alpha, model = "Gaussian", weights = "equal", penalize.diagonal = FALSE)
```

Arguments

X	A matrix ($K \times G$) of list of data matrices ($n_g \times p$), where K is the number of data types and G is the number of subpopulations. The (k,g)-th element is a $n_g \times p$ data matrix corresponding to the k-th data type and the g-th subpopulation.
lambda	The tuning parameter for controlling the level of sparsity of networks.
alpha	The tuning parameter for controlling the extend of similarity among subpopulation-specific networks.
model	A character string indicating which model is used to fit the data. "Gaussian" (default) and "nonparanormal" can be used. If model == "Gaussian", the Gaussian graphical model will be used. If model == "nonparanormal", the nonparanormal graphical model will be used.
weights	Determines the putative sample size of each subpopulation's data. Allowed values: a vector with length equal to the number of subpopulations; "equal", giving each subpopulation weight 1; "sample.size", giving each subpopulation weight corresponding to its sample size.
penalize.diagonal	Determines whether the sparsity penalty is applied to the diagonal.

Details

The function is used to infer gene networks from gene expression data.

Value

Omega.hat	A matrix ($K \times G$) of list of estimated precision matrices.
R.hat	A matrix ($K \times 1$) of list of estimated common components.
M.hat	A matrix ($K \times G$) of list of estimated subpopulation-unique components.
Omega.bar	A matrix ($1 \times G$) of list of estimated gene networks.
R.bar	A matrix of estimated common subnetworks.
M.bar	A matrix ($1 \times G$) of list of estimated subpopulation-unique subnetworks.

Author(s)

Xiao-Fei Zhang

References

Xiao-Fei Zhang, Le Ou-Yang, Shuo Yang, Hong Qin, Xiaohua Hu and Hong Yan (2017), JGNI: A graphical model for joint gene network inference across multiple subpopulations and data types.

See Also

admm, generate_data, Corr, TCGA.BRCA

Examples

```
# Simulation data
dat = generate_data(K = 3, G = 4, n = 100, p = 100, tau = 0.5, network.type = "ER")
result = JGNI(dat$X, 0.6, 0.3, model = "Gaussian", weights = "equal", penalize.diagonal = FALSE)

# TCGA breast cancer data
data("TCGA.BRCA")
result = JGNI(TCGA.BRCA$X, 1.1, 0.4, model = "nonparanormal", weights = "equal")
```

TCGA.BRCA *TCGA breast cancer data*

Description

The TCGA breast cancer gene expression data used in our study. The data are obtained from the TCGA database. They are collected from three platforms: mRNA expression (Agilent G450 microarray), mRNA expression (RNA sequencing), and copy number variants (Affymetrix genome-wide human SNP Array 6.0). The cancers are grouped into four subtypes: Luminal A, Luminal B, HER2-enriched, and Basal-like. The data only include expression measure of genes that overlap with the breast cancer pathway collected from the Kyoto Encyclopedia of Genes and Genomes database.

Usage

```
data("TCGA.BRCA")
```

Author(s)

Xiao-Fei Zhang

Source

The Cancer Genome Atlas Research Network (2012), Comprehensive molecular portraits of human breast tumors. Nature. 490(7418), 61-70. (http://cancergenome.nih.gov/)

References

Xiao-Fei Zhang, Le Ou-Yang, Shuo Yang, Hong Qin, Xiaohua Hu and Hong Yan (2017), JGNI: A graphical model for joint gene network inference across multiple subpopulations and data types.

See Also

JGNI, admm, Corr

Examples

```
data(TCGA.BRCA)
## maybe str(TCGA.BRCA) ...
```

Index