Package 'mapggm'

October 7, 2015

Type Package

Title Multi-attribute network construction and perturbation detection

Version 1.0

Date 2015-09-30

Author Paula J Griffin

Maintainer Paula J Griffin <paulajgriffin@gmail.com>

Description
    Companion package to 'Detection of multiple perturbations in multi-omics biological networks'

License Apache License

Imports igraph, Matrix

R topics documented:

---

findPerturbations   *Get node-wise test statistics*

Description

    Given a vector `id` indicating node membership, test for perturbations at each node as described by
    Griffin et al. This function performs either the single-target test (`sequential=FALSE`, default) or
    the multi-target adjusted procedure (`sequential=TRUE`)

Usage

```
findPerturbations(Y, Omega, Sigma, id, sequential = FALSE)
```

Arguments

`Y`	matrix or data frame (rows=subjects, columns=variables)
`Omega`	precision matrix
`Sigma`	covariance matrix
`id`	vector mapping variables to nodes
`sequential`	boolean, whether or not to perform sequential adjustments

Value

test statistics for perturbations at each individual node

References

Griffin, P.J., Johnson, W.E., and Kolaczyk, E.K. Detection of multiple perturbations in multi-omics biological networks' (under review)

`getLambdas`	*Generate lambdas to try*

Description

Generate penalty parameters spanning a range of 1 to `length(unique(id))` submatrix blocks. This function is provided to establish a range of "reasonable" penalty parameters for optimization according to the algorithm of Kolar et al (2014).

Usage

```
getLambdas(S, id, length.out = 10)
```

Arguments

`S`	sample covariance
`id`	vector of node identifiers
`length.out`	number of lambda parameters to return

Value

vector of `length.out` equally spaced lambdas

References

Kolar, M., Liu, H., and Xing, E. P. (2014). Graph estimation from multi-attribute data. The Journal of Machine Learning Research 15, 1713<e2><80><93>1750.

Examples

```
Y <- matrix(rnorm(120), nrow=20, ncol=6)
S <- crossprod(Y)
id <- rep(1:3, each=2)
getLambdas(S, id, 5)
```

---

groupCondLRTs *Sequentially test for perturbations at a series of locations*

Description

Given a matrix of possible perturbation locations, perform the sequential testing procedure described by Griffin et al. Return either test statistics or p-values.

Usage

```
groupCondLRTs(Y, Omega, Sigma, perturb.mat, ret = "stat")
```

Arguments

Y	matrix or data frame (rows=subjects, columns=variables)
Omega	precision matrix
Sigma	covariance matrix
perturb.mat	matrix indicating sets of which variables to test for perturbations
ret	return statistics ('stat') or p-values ('pval')

Value

test statistics for perturbations at each of perturb.mat rows

References

Griffin, P.J., Johnson, W.E., and Kolaczyk, E.K. Detection of multiple perturbations in multi-omics biological networks' (under review)

---

groupLRTs *Test for perturbations at a series of locations*

Description

Given a matrix of possible perturbation locations, perform the single-target (non-sequential) testing procedure described by Griffin et al. Return either test statistics or p-values.

Usage

```
groupLRTs(Y, Omega, Sigma, perturb.mat, ret = "stat")
```

Arguments

Y	matrix or data frame (rows=subjects, columns=variables)
Omega	precision matrix
Sigma	covariance matrix
perturb.mat	matrix indicating sets of which variables to test for perturbations
ret	value to return ('stat' or 'pval')

Value

test statistics for perturbations at each of perturb.mat rows

References

Griffin, P.J., Johnson, W.E., and Kolaczyk, E.K. Detection of multiple perturbations in multi-omics biological networks' (under review)

---

maLasso   *Multi-attribute network estimation*

Description

Estimates block precisions and covariances for a multi-attribute network based on a Gaussian graphical model with zero mean vector. This is according to the method described by Kolar et al (2015).

Usage

```
maLasso(S, n, id, lambda, W = NULL, update = 100, max.gap = 0.5,
  max.iter = 100, min.t = .Machine$double.eps)
```

Arguments

S	sample covariance matrix
n	number of samples
id	vector of node identifiers
lambda	tuning parameter for penalty
W	optional penalty weight matrix (same dimensions as S)
update	how often to print updates
max.gap	maximum allowable primal-dual gap
max.iter	maximum number of iterations to complete (overriedes max.gap)
min.t	minimum step size (overrides max.gap, max.iter)

Value

list of precision, covariance, optimization status, lambda, and number components

References

Kolar, M., Liu, H., and Xing, E. P. (2014). Graph estimation from multi-attribute data. The Journal of Machine Learning Research 15, 1713<e2><80><93>1750.

---

runLasso | *Wrapper function for multi-attribute network estimation via lasso*

Description

Performs estimation of a zero-mean multi-attribute Gaussian graphical model as described by Kolar et al (2014) when mode=1.

Usage

```
runLasso(S, n, id, lambda, W = NULL, mode = 1, update = 100,
  max.gap = 0.5, max.iter = 100, min.t = .Machine$double.eps)
```

Arguments

S	sample covariance matrix
n	number of samples
id	vector assigning variables to nodes
lambda	penalty tuning parameter
W	optional weight matrix
mode	estimation mode. Can be 1 (multiattribute; default), 2 (unstructured together), or 3 (separately; only valid for 2-attribute data)
update	how often to print updates in optimization
max.gap	maximum allowable primal/dual gap
max.iter	maximum number of iterations to optimize (overrides max.gap)
min.t	minimum step size (overrides max.gap, max.iter)

Details

If mode=2, the id argument is essentially disregarded, and all rows/columns in S are treated as if they represent individual nodes. If mode=3, the optimization of mode=2 is performed separately for each attribute type, and a matrix with 0 on all cross-attribute entries is returned.

Value

list of precision, covariance, optimization status, lambda, and number components

References

Kolar, M., Liu, H., and Xing, E. P. (2014). Graph estimation from multi-attribute data. The Journal of Machine Learning Research 15, 1713<e2><80><93>1750.

---

runLassoSelect   *Wrapper function for multi-attribute network estimation plus selection
                                via lasso*

Description

Performs estimation of a zero-mean multi-attribute Gaussian graphical model as described by Kolar
et al (2014) when mode=1 with a selection procedure to determine optimal lambda. Selection may
be performed according to extended BIC, BIC, AIC, or Hamming distance to true graph. The true
graph Theta is only required if Hamming distance is being used for the selection procedure.

Usage

```
runLassoSelect(S, n, id, lambda.range, W = NULL, mode = 1, update = 100,
  max.gap = 0.5, max.iter = 100, min.t = .Machine$double.eps,
  method = "EBIC", Theta.true = NULL, plot = NULL, gamma = 0.5)
```

Arguments

S	sample covariance matrix
n	number of samples
id	vector assigning variables to nodes
lambda.range	vector of lambdas to try
W	optional weight matrix
mode	estimation mode. Can be 1 (multiattribute; default), 2 (unstructured together), or 3 (separately; only valid for 2-attribute data)
update	how often to print updates in optimization
max.gap	maximum allowable primal/dual gap
max.iter	maximum number of iterations to optimize (overrides max.gap)
min.t	minimum step size (overrides max.gap, max.iter
method	how to select the best model ('EBIC', 'BIC', 'AIC', or 'hamming')
Theta.true	true underlying graph (needed for 'hamming' method only)
plot	boolean, whether or not to make diagnostic plot
gamma	gamma parameter for EBIC method (default 0.5)

Details

If mode=2, the id argument is essentially disregarded, and all rows/columns in S are treated as if
they represent individual nodes. If mode=3, the optimization of mode=2 is performed separately for
each attribute type, and a matrix with 0 on all cross-attribute entries is returned.

Value

list of precision, covariance, optimization status, lambda, and number components

References

Kolar, M., Liu, H., and Xing, E. P. (2014). Graph estimation from multi-attribute data. The Journal of Machine Learning Research 15, 1713<e2><80><93>1750.

Chen, J. and Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. Biometrika 95, 759<e2><80><93>771.

Index