Package ‘spaceExt’

February 20, 2015

Version 1.0
Date 2011-11-10
Title Extension of SPACE
Author Shiyuan He <flyoverearth@163.com>
Maintainer Shiyuan He <flyoverearth@163.com>
Depends limSolve, glasso
Description undirected graph inference with missing data
License GPL (>= 2)
Repository CRAN
Date/Publication 2011-12-16 18:26:25
NeedsCompilation yes

R topics documented:

finStocksCn ............................................................. 1
glasso.miss ............................................................ 2
space.miss .............................................................. 3

Index

finStocksCn Continuously Compounded Daily Returns for Financial Sector of Chinese Stock Market

Description
This data set contains the continuously compounded daily returns for financial sector of Chinese stock market. The time ranges from January 4th, 2011 to July 27th, 2011. It totally consists of 34 stocks: 3 insurance stocks, 15 security stocks and 16 bank stocks.

Usage

finStocksCn
Format

A list containing two components

1. $\text{Names}$, a vector of length 34 containing the stock names.
2. $\text{Returns}$, a 155-by-34 matrix, each column corresponding to a stock and each row to a trading day. The stock name of its $i^{(th)}$ column is contained in the $i^{(th)}$ component of $\text{StocksCn}$. $\text{Names}$.

Description

A function to achieve covariance selection by glasso with EM, missing data is allowed

Usage

glasso.miss(\text{Y, emIter, rho, ...})

Arguments

\text{Y} \quad \text{numeric matrix. Columns are for variables and rows are for samples.}
\text{emIter} \quad \text{numeric value. The maximum number of iteration for EM}
\text{rho} \quad \text{The l1 norm tuning parameter.}
\text{...} \quad \text{additional control parameter passed to glasso}

Value

a list of following components.

\text{Y.imputed} \quad \text{The sample matrix Y with missing data imputed by EM}
\text{bic} \quad \text{the BIC(Bayesian Information Criterion) of current fit.}
\text{...} \quad \text{other values returned by glasso}

Author(s)

Shiyuan He

References

Examples

```r
data(finstockscn) # data of finance sector of Chinese Stock Market, from January 4th, 2011 to July 26th, 2011
finstockscn$names # stock names
y.m<-scale(finstockscn$returns)
n=nrow(y.m)
p=ncol(y.m)
res=glasso.miss(y.m,rho=0.03,emIter=25,penalize.diagonal=FALSE)
res$bic # bic returned
```

---

**space.miss**

*Sparse Covariance Selection by SPACE with EM*

**Description**

A function to estimate partial correlations using SPACE method with EM, missing data is allowed

**Usage**

```r
space.miss(Y.m, lam1, lam2=0, sig=NULL, weight=NULL, iter=2, emIter=5, n_iter=1000, t0=0, r=0)
```

**Arguments**

- **Y.m**
  numeric matrix. Columns are for variables and rows are for samples. It’s recommended to first standardize each column to have mean 0 and $l_2$ norm 1.

- **lam1**
  numeric value. This is the $l_1$ norm penalty parameter. If the columns of Y.m have norm one, then the suggested range of lam1 is $O(n^{3/2}\Phi^{-1}(1 - \alpha/(2p^2)))$ for small $\alpha$ such as 0.1.

- **lam2**
  numeric value. If not specified, lasso regression is used in the Joint Sparse Regression Model (JSRM). Otherwise, elastic net regression is used in JSRM and lam2 serves as the $l_2$ norm penalty parameter.

- **sig**
  numeric vector. Its length should be the same as the number of columns of Y.m. It is the vector of $\sigma^{ii}$ (the diagonal of the inverse covariance matrix). If not specified, $\sigma^{ii}$ will be estimated during the model fitting with initial values rep(1,p). The number of the iteration of the model fitting (iter) will then be at least 2. Note, the scale of sig does not matter.

- **weight**
  numeric value or vector. It specifies the weights or the type of weights used for each regression in JSRM. The default value is NULL, which means all regressions will be weighted equally in the joint model. If weight=1, residue variances will be used for weights. If weight=2, the estimated degree of each variable will be used for weights. Otherwise, it should be a positive numeric vector, whose length is equal to the number of columns of Y.m

- **iter**
  integer. It is the total number of interactions in JSRM for estimating $\sigma^{ii}$ and partial correlations. When sig=NULL and/or weight=NULL or 2, iter should be at least 2.

- **emIter**
  integer. the maximum number of EM iteration allowed
space.miss

n_iter integer. the maximum number of interations in JSRM.

t0 integer. 1 <= t0 <= n. the time point at which to perform local smoothing estimation.

r positive value. local smoothing parameter. If r=0, then no smoothing is incorporated.

Details

space.miss Based on the work of J. Peng, et al(2007), this function allows computing with missing data by implementing EM algorithm.

\[
\frac{1}{2} \sum_{i=1}^{n} K(r \cdot |t - t_0|)[\sum_{i=1}^{p} w_i(y_{ti} - \sum_{j \neq i} \sqrt{\frac{\sigma_{jj}}{\sigma_{ii}}} \rho_{ij} y_{tj})^2] + \lambda \sum_{i<j} |\rho_{ij}|
\]

where \(K(\cdot)\) is the smoothing kernal, \(K(x) = \exp(-x)\).

Value

a list of following components.

- Y.imputed The sample matrix \(Y\) with missing data imputed by EM
- ParCor the estimated partial correlation matrix.
- sig.fit numeric vector of the estimated diagonal \(\sigma_{ii}\).
- bic BIC for the current estimate.

Author(s)

Shiyuan He

References


Examples

data(finstockscn) #data of finance sector of Chinese Stock Market, from January 4th, 2011 to July 26th, 2011
finStocksCn$names #stock names
y.m<-scale(finstocksCn$returns)
n=nrow(y.m)
p=ncol(y.m)
alpha=0.1
l1=1/sqrt(n)*qnorm(1-alpha/(2*p^2))
res=space.miss(y.m,lambda=l1*25,emIter=25,iter=2,weight=2)
res$bic #bic returned
Index

*Topic datasets
  finStocksCn, 1
*Topic methods
  glasso, 2
  glasso.miss, 2
  space.miss, 3

finStocksCn, 1
glasso, 2
glasso.miss, 2
space.miss, 3