Package ‘ShrinkCovMat’

July 30, 2019

Type Package
Title Shrinkage Covariance Matrix Estimators
Version 1.4.0
Description Provides nonparametric Steinian shrinkage estimators of the covariance matrix that are suitable in high dimensional settings, that is when the number of variables is larger than the sample size.
License GPL-2 | GPL-3
Depends R (>= 2.10)
LazyData true
NeedsCompilation yes
Encoding UTF-8
URL http://github.com/AnestisTouloumis/ShrinkCovMat
BugReports http://github.com/AnestisTouloumis/ShrinkCovMat/issues
RoxygenNote 6.1.1
Suggests testthat (>= 2.1.0), covr
Imports Rcpp (>= 1.0.1)
LinkingTo Rcpp, RcppArmadillo
Author Anestis Touloumis [aut, cre] (0000-0002-5965-1639)
Maintainer Anestis Touloumis <A.Touloumis@brighton.ac.uk>
Repository CRAN
Date/Publication 2019-07-30 08:30:02 UTC

R topics documented:

ShrinkCovMat-package .................................................. 2
colon ................................................................. 3
shrinkcovmat.equal ............................................... 3
shrinkcovmat.identity ........................................... 5
shrinkcovmat.unequal ........................................... 6
targetselection .................................................... 7
**Description**

Provides nonparametric Stein-type shrinkage estimators of the covariance matrix that are suitable and statistically efficient when the number of variables is larger than the sample size. These estimators are non-singular and well-conditioned regardless of the dimensionality.

**Details**

Each of the implemented shrinkage covariance matrix estimators is a convex linear combination of the sample covariance matrix and of a target matrix. Three options are considered for the target matrix: (a) the diagonal matrix with diagonal elements the average of the sample variances (\texttt{shrinkcovmat.equal}), (b) the diagonal matrix with diagonal elements the corresponding sample variances (\texttt{shrinkcovmat.unequal}), and (c) the identity matrix (\texttt{shrinkcovmat.identity}). The optimal shrinkage intensity determines how much the sample covariance matrix will be shrunk towards the selected target matrix. Estimation of the corresponding optimal shrinkage intensities is discussed in Touloumis (2015). The function \texttt{targetselection} is designed to ease the selection of the target matrix.

**Author(s)**

Anestis Touloumis

Maintainer: Anestis Touloumis <A.Touloumis@brighton.ac.uk>

**References**


**See Also**

Useful links:

- [http://github.com/AnestisTouloumis/ShrinkCovMat](http://github.com/AnestisTouloumis/ShrinkCovMat)
- Report bugs at [http://github.com/AnestisTouloumis/ShrinkCovMat/issues](http://github.com/AnestisTouloumis/ShrinkCovMat/issues)
Colon Cancer Dataset

Description

The dataset describes a colon cancer study (Alon et al., 1999) in which gene expression levels were measured on 40 normal tissues and on 22 tumor colon tissues. Note that a logarithmic (base 10) transformation has been applied to the gene expression levels.

Format

A data frame in which the rows correspond to 2000 genes and the columns to 62 tissues. The first 40 columns belong to the normal tissue group while the last 22 columns to the tumor colon tissue group.

Source

http://genomics-pubs.princeton.edu/oncology/affydata

References


Examples

data(colon)
summary(colon)

Shrink the Sample Covariance Matrix Towards a Sphericity Matrix

Description

Provides a nonparametric Stein-type shrinkage estimator of the covariance matrix that is a linear combination of the sample covariance matrix and of a diagonal matrix with the average of the sample variances on the diagonal and zeros elsewhere.

Usage

shrinkcovmat.equal(data, centered = FALSE)

1http://genomics-pubs.princeton.edu/oncology/affydata
Arguments

data  a numeric matrix containing the data.
centered  a logical indicating if the mean vector is the zero vector.

Details

The rows of the data matrix data correspond to variables and the columns to subjects.

Value

Returns an object of the class 'shrinkcovmathat' that has components:

SigmaHat  The Stein-type shrinkage estimator of the covariance matrix.
lambdahat  The estimated optimal shrinkage intensity.
sigmasample  The sample covariance matrix.
Target  The target covariance matrix.
centered  If the data are centered around their mean vector.

Author(s)

Anestis Touloumis

References


See Also

shrinkcovmat.unequal and shrinkcovmat.identity.

Examples

data(colon)
normal_group <- colon[, 1:40]
tumor_group <- colon[, 41:62]
sigma_hat_normal_group <- shrinkcovmat.equal(normal_group)
sigma_hat_normal_group
sigma_hat_tumor_group <- shrinkcovmat.equal(tumor_group)
sigma_hat_tumor_group
Shrinking the Sample Covariance Matrix Towards the Identity Matrix

Description

Provides a nonparametric Stein-type shrinkage estimator of the covariance matrix that is a linear combination of the sample covariance matrix and of the identity matrix.

Usage

shrinkcovmat.identity(data, centered = FALSE)

Arguments

data

a numeric matrix containing the data.

centered

a logical indicating if the mean vector is the zero vector.

Details

The rows of the data matrix data correspond to variables and the columns to subjects.

Value

Returns an object of the class 'shrinkcovmathat' that has components:

Sigmahat

The Stein-type shrinkage estimator of the covariance matrix.

lambdahat

The estimated optimal shrinkage intensity.

Sigmasample

The sample covariance matrix.

Target

The target covariance matrix.

centered

If the data are centered around their mean vector.

Author(s)

Anestis Touloumis

References


See Also

shrinkcovmat.equal and shrinkcovmat.unequal.
Examples

```r
data(colon)
normal_group <- colon[, 1:40]
tumor_group <- colon[, 41:62]
sigma_hat_normal_group <- shrinkcovmat.identity(normal_group)
sigma_hat_normal_group
sigma_hat_tumor_group <- shrinkcovmat.identity(tumor_group)
sigma_hat_tumor_group
```

shrinkcovmat.unequal

*Shrinking the Sample Covariance Matrix Towards a Diagonal Matrix with Diagonal Elements the Sample Variances.*

Description

Provides a nonparametric Stein-type shrinkage estimator of the covariance matrix that is a linear combination of the sample covariance matrix and of the diagonal matrix with elements the corresponding sample variances on the diagonal and zeros elsewhere.

Usage

```r
shrinkcovmat.unequal(data, centered = FALSE)
```

Arguments

- **data**: a numeric matrix containing the data.
- **centered**: a logical indicating if the vectors are centered around their mean vector.

Details

The rows of the data matrix `data` correspond to variables and the columns to subjects.

Value

Returns an object of the class 'shrinkcovmathat' that has components:

- **Sigmahat**: The Stein-type shrinkage estimator of the covariance matrix.
- **lambdahat**: The estimated optimal shrinkage intensity.
- **Sigmasample**: The sample covariance matrix.
- **Target**: The target covariance matrix.
- **centered**: If the data are centered around their mean vector.

Author(s)

Anestis Touloumis
**targetselection**

**Target Matrix Selection**

**Description**

Implements the rule of thumb proposed by Touloumis (2015) for target matrix selection. If the estimated optimal shrinkage intensities of the three target matrices are of similar magnitude, then the average and the range of the sample variances should be inspected in order to adopt the most plausible target matrix.

**Usage**

```r
targetselection(data, centered = FALSE)
```

**Arguments**

- `data`: a numeric matrix containing the data.
- `centered`: a logical indicating if the mean vector is the zero vector.

**Details**

The rows of the data matrix `data` correspond to variables and the columns to subjects.
Value

Prints the estimated optimal shrinkage intensities, the range and average of the sample variances and returns an object of the class 'targetsel' that has components:

- `optimal_sphericity`  
The estimated optimal intensity for a target matrix with equal variances.
- `optimal_identity`  
The estimated optimal shrinkage intensity for the identity target matrix.
- `optimal_diagonal`  
The estimated optimal intensity for a target matrix with unequal variances.
- `range`  
The range of the sample variances.
- `average`  
The average of the sample variances.

Author(s)

Anestis Touloumis

References


Examples

data(colon)
normal_group <- colon[, 1:40]
targetselection(normal_group)
## Similar intensities, the range of the sample variances is small and the
## average is not close to one. The scaled identity matrix seems to be the
## most suitable target matrix for the normal group.

tumor_group <- colon[, 41:62]
targetselection(tumor_group)
## Similar intensities, the range of the sample variances is small and the
## average is not close to one. The scaled identity matrix seems to be the
## most suitable target matrix for the colon group.