Package: inet (via r-universe)

September 11, 2024

Type Package

Title Performing Inference on Networks with Regularization

Version 0.1.0

Depends R (>= 3.5.0)

Author Lourens Waldorp <waldorp@uva.nl>, Jonas Haslbeck <jonashaslbeck@gmail.com>

Maintainer Jonas Haslbeck <jonashaslbeck@gmail.com>

Description Performs inference with the lasso in Gaussian Graphical Models. The package consists of wrappers for functions from the hdi package.

Encoding UTF-8

LazyData true

License GPL (>= 2)

Imports hdi, glmnet

Repository https://jmbh.r-universe.dev

RemoteUrl https://github.com/jmbh/inet

RemoteRef HEAD

RemoteSha d1cd82eb065db7d95f9880c051fb78c4270411d4

Contents

inet-datasets *Datasets included in inet package*

Description

The package includes a dataset with measurements of 17 PTSD symptoms taken from 344 individuals. See McNally et al. (2015) for more details.

Author(s)

Jonas Haslbeck

References

McNally, R. J., Robinaugh, D. J., Wu, G. W., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to posttraumatic stress disorder. Clinical Psychological Science, 3(6), 836-849.

inet-internal *Internal inet functions*

Description

Internal inet functions.

Details

These are internal functions.

Author(s)

Jonas Haslbeck

`lasso` *Estimate GGM with nodewise regression and the lasso.*

Description

Estimate a Gaussian Graphical Model with lasso-regularized nodewise regression, where the regularization parameter is selected with cross-validation. This is a wrapper around the function `cv.glmnet()` from the `glmnet` package.

Usage

```
lasso(data, pbar = TRUE, nfolds = 10, rulereg = "and")
```

Arguments

`data`	An n x p matrix containing the data, where n are cases and p are variables
`pbar`	If `pbar = TRUE`, a progress bar will be displayed.
`nfolds`	Specifies the number of folds used to select the regularization parameter in each of the p nodewise regressions.
`rulereg`	Specifies how parameter estimates should be combined across nodewise regressions. The options are the AND-rule (requiring both estimates to be significant) or the OR-rule (only requiring one estimate to be significant). Defaults to `rulereg = "and"`.

Value

The function returns a list with the following entries:

`est`	A p x p matrix with point estimates for all partial correlations
`select`	A p x p indicator matrix indicating which edges have been selected to be present.
`ints`	A p-vector of estimated intercepts.

Author(s)

Jonas Haslbeck <jonashaslbeck@gmail.com>

References

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of statistical software, 33(1), 1.

Examples

```
## Not run:

# Fit GGM to PTSD data
set.seed(1)
out <- lasso(data = ptsd_data)


## End(Not run)
```

`lasso_dsp` *Estimate GGMs with the desparsified lasso.*

Description

Estimate Gaussian Graphical Models using the desparsified lasso. This is a wrapper around the function `lasso.proj` of the `hdi` package.

Usage

```
lasso_dsp(data, betainit = "cv lasso", ci.level = 0.95,
          correction = TRUE, pbar = TRUE, rulereg = "and")
```

Arguments

`data`	An n x p matrix containing the data, where n are cases and p are variables
`betainit`	Specifying how to estimate lasso solution in initial estimation. Either `betainit = "cv lasso"` (default) or `betainit = "cv lasso"`. See the manual of the function `lasso.proj` of the `hdi` package for more info.
`ci.level`	Specifies the width of the confidence interval used for testing the null hypothesis that a parameter is different to zero. Defaults to `ci.level = 0.95`, which corresponds to a critical threshold of 0.05.
`correction`	If `correction = TRUE`, the Bonferroni-Holm correction will be applied to p-values on the level of nodewise regressions (see e.g., Hochberg, 1987).
`pbar`	If `pbar = TRUE`, a progress bar will be displayed.
`rulereg`	Specifies how parameter estimates should be combined across nodewise regressions. The options are the AND-rule (requiring both estimates to be significant) or the OR-rule (only requiring one estimate to be significant). Defaults to `rulereg = "and"`.

Value

The function returns a list with the following entries:

`est`	A p x p matrix with point estimates for all partial correlations
`est.signf`	A p x p matrix with point estimates for all partial correlations with non-significant partial correlations being thresholded to zero.
`signf`	A p x p matrix indicating for each partial correlation whether it is significantly different to zero.
`ci.lower`	A p x p matrix indicating the lower confidence interval for each partial correlation.
`ci.upper`	A p x p matrix indicating the upper confidence interval for each partial correlation.

Author(s)

Jonas Haslbeck <jonashaslbeck@gmail.com>; Lourens Waldorp <waldorp@uva.nl>

References

Hochberg, Y., & Tamhane, A. C. (1987). Multiple comparison procedures. John Wiley & Sons, Inc..

Buehlmann, P., Kalisch, M., & Meier, L. (2014). High-dimensional statistics with a view toward applications in biology. Annual Review of Statistics and Its Application, 1, 255-278.

Examples

```
## Not run:

# Fit GGM to PTSD data
set.seed(1)
out <- lasso_dsp(data = ptsd_data)


## End(Not run)
```

`lasso_dsp_boot` *Estimate GGMs with the desparsified lasso using the bootstrap.*

Description

Estimate Gaussian Graphical Models using the desparsified lasso using the bootstrap. This is a wrapper around the function `lasso.proj` of the `hdi` package.

Usage

```
lasso_dsp_boot(data, betainit = "cv lasso", ci.level = 0.95,
               correction = TRUE, B = 1000, pbar = TRUE,
               rulereg = "and")
```

Arguments

data	An n x p matrix containing the data, where n are cases and p are variables
betainit	Specifying how to estimate lasso solution in initial estimation. Either betainit = "cv lasso" (default) or betainit = "cv lasso". See the manual of the function lasso.proj of the hdi package for more info.
ci.level	Specifies the width of the confidence interval used for testing the null hypothesis that a parameter is different to zero. Defaults to ci.level = 0.95, which corresponds to a critical threshold of 0.05.
correction	If correction = TRUE, the Bonferroni-Holm correction will be applied to p-values on the level of nodewise regressions (see e.g., Hochberg, 1987).
B	The number of bootstrap samples used for estimation. Defaults to B=1000.
pbar	If pbar = TRUE, a progress bar will be displayed.
rulereg	Specifies how parameter estimates should be combined across nodewise regressions. The options are the AND-rule (requiring both estimates to be significant) or the OR-rule (only requiring one estimate to be significant). Defaults to rulereg = "and".

Value

The function returns a list with the following entries:

est	A p x p matrix with point estimates for all partial correlations
est.signf	A p x p matrix with point estimates for all partial correlations with non-significant partial correlations being thresholded to zero.
signf	A p x p matrix indicating for each partial correlation whether it is significantly different to zero.
ci.lower	A p x p matrix indicating the lower confidence interval for each partial correlation.
ci.upper	A p x p matrix indicating the upper confidence interval for each partial correlation.

Author(s)

Jonas Haslbeck <jonashaslbeck@gmail.com>; Lourens Waldorp <waldorp@uva.nl>

References

Hochberg, Y., & Tamhane, A. C. (1987). Multiple comparison procedures. John Wiley & Sons, Inc..

Bühlmann, P., Kalisch, M., & Meier, L. (2014). High-dimensional statistics with a view toward applications in biology. Annual Review of Statistics and Its Application, 1, 255-278.

Davison, A. C., & Hinkley, D. V. (1997). Bootstrap methods and their application (No. 1). Cambridge university press.

Examples

```
## Not run:

# Fit GGM to PTSD data
set.seed(1)
out <- lasso_dsp_boot(data = ptsd_data)


## End(Not run)
```

`lasso_ms` *Estimate GMM with inference via the multi-split method.*

Description

Estimate Gaussian Graphical Models with inference base don the multi-split method. This is a wrapper of the function `multi.split` of the `hdi` package.

Usage

```
lasso_ms(data, B = 50, fraction = 0.5, ci.level = 0.95,
         correction = TRUE, pbar = TRUE, rulereg = "and")
```

Arguments

`data`	An n x p matrix containing the data, where n are cases and p are variables
`B`	The number of sample-splits. Defaults to `B=50`.
`fraction`	a number in (0,1), the fraction of data used at each sample split for the model selection process. The remaining data is used for calculating the p-values.
`ci.level`	Specifies the width of the confidence interval used for testing the null hypothesis that a parameter is different to zero. Defaults to `ci.level = 0.95`, which corresponds to a critical threshold of 0.05.
`correction`	If `correction = TRUE`, the Bonferroni-Holm correction will be applied to p-values on the level of nodewise regressions (see e.g., Hochberg, 1987).
`pbar`	If `pbar = TRUE`, a progress bar will be displayed.

rulereg	Specifies how parameter estimates should be combined across nodewise regressions. The options are the AND-rule (requiring both estimates to be significant) or the OR-rule (only requiring one estimate to be significant). Defaults to rulereg = "and".

Value

The function returns a list with the following entries:

est	A p x p matrix with point estimates for all partial correlations
est.signf	A p x p matrix with point estimates for all partial correlations with non-significant partial correlations being thresholded to zero.
signf	A p x p matrix indicating for each partial correlation whether it is significantly different to zero.
ci.lower	A p x p matrix indicating the lower confidence interval for each partial correlation.
ci.upper	A p x p matrix indicating the upper confidence interval for each partial correlation.

Author(s)

Jonas Haslbeck <jonashaslbeck@gmail.com>; Lourens Waldorp <waldorp@uva.nl>

References

Hochberg, Y., & Tamhane, A. C. (1987). Multiple comparison procedures. John Wiley & Sons, Inc..

Wasserman, L., & Roeder, K. (2009). High dimensional variable selection. Annals of statistics, 37(5A), 2178.

Meinshausen, N., Meier, L., & Bühlmann, P. (2009). P-values for high-dimensional regression. Journal of the American Statistical Association, 104(488), 1671-1681.

Examples

```
## Not run:

# Fit GGM to PTSD data
set.seed(1)
out <- lasso_ms(data = ptsd_data)


## End(Not run)
```

OLS *Estimate GMM via nodewise regression and hypothesis tests.*

Description

Estimate Gaussian Graphical Model with nodewise regression, selecting edges with standard hypothesis tests and the Bonferroni-Holm Correction.

Usage

```
OLS(data, pbar = TRUE, correction = TRUE,
    ci.level = 0.95, rulereg = "and")
```

Arguments

`data`	An n x p matrix containing the data, where n are cases and p are variables
`pbar`	If `pbar = TRUE`, a progress bar will be displayed.
`correction`	If `correction = TRUE`, the Bonferroni-Holm correction will be applied to p-values on the level of nodewise regressions (see e.g., Hochberg, 1987).
`ci.level`	Specifies the width of the confidence interval used for testing the null hypothesis that a parameter is different to zero. Defaults to `ci.level = 0.95`, which corresponds to a critical threshold of 0.05.
`rulereg`	Specifies how parameter estimates should be combined across nodewise regressions. The options are the AND-rule (requiring both estimates to be significant) or the OR-rule (only requiring one estimate to be significant). Defaults to `rulereg = "and"`.

Value

The function returns a list with the following entries:

`est`	A p x p matrix with point estimates for all partial correlations
`est.signf`	A p x p matrix with point estimates for all partial correlations with non-significant partial correlations being thresholded to zero.
`signf`	A p x p matrix indicating for each partial correlation whether it is significantly different to zero.
`ci.lower`	A p x p matrix indicating the lower confidence interval for each partial correlation.
`ci.upper`	A p x p matrix indicating the upper confidence interval for each partial correlation.
`ints`	A p-vector of estimated intercepts.

Author(s)

Jonas Haslbeck <jonashaslbeck@gmail.com>

References

Hochberg, Y., & Tamhane, A. C. (1987). Multiple comparison procedures. John Wiley & Sons, Inc..

Examples

```
## Not run:

# Fit GGM to PTSD data
out <- OLS(data = ptsd_data)


## End(Not run)
```

`plot.inet` *Plot point estimates and confidence intervals*

Description

Plot point estimates and confidence intervals for models estimated with the `lasso_ms`, `lasso_dsp`, `lasso_dsp_boot` and `OLS` functions.

Usage

```
## S3 method for class 'inet'
plot(x, labels = NULL, order = FALSE, subset = NULL,
          cex.labels = 0.80, cex.axis = 0.75, ...)
```

Arguments

`x`	The output object from either `lasso_ms`, `lasso_dsp`, `lasso_dsp_boot` or `OLS`.
`labels`	A p-vector of characters specifying the labels for variables.
`order`	If `order = TRUE`, the edges are listed in decreasing order based on the point estimate.
`subset`	Allows to only display a subset of the edges. For example, if `subset=1:20` the first 20 edges are displayed. This is especially useful for larger networks, in which all edges are unlikely to fit into a single figure.
`cex.labels`	The font size of the edge labels.
`cex.axis`	The font size of the axes.
`...`	Additional arguments.

Author(s)

Jonas Haslbeck <jonashaslbeck@gmail.com>

Examples

```
## Not run:

# Fit GGM to PTSD data
set.seed(1)
out <- lasso_dsp(data = ptsd_data)

# Plot first 20 edges
plot(out, labels = colnames(ptsd_data),
     order=TRUE, subset = 1:20)


## End(Not run)
```

Index