Title  Exploratory Graph Analysis – a Framework for Estimating the Number of Dimensions in Multivariate Data using Network Psychometrics

Version 1.1.0

Date 2022-05-09

Maintainer Hudson Golino <hfg9s@virginia.edu>


Depends R (>= 3.5.0)

License GPL (>= 3.0)

Encoding UTF-8

LazyData true

Imports qgraph (>= 1.4.1), semPlot (>= 1.0.1), igraph (>= 1.0.1), lavaan (>= 0.5-22), glasso (>= 1.10), stats, Matrix (>= 1.2),
network (>= 1.16.0), GGally (>= 2.0.0), methods, fitdistrplus, ggpubr, ggplot2, pbapply, matrixcalc (>= 1.0-3), OpenMx

Suggests psych, psychTools, GPArotation, sna, markdown, knitr, markdown, gridExtra, rstudioapi, RColorBrewer

VignetteBuilder knitr
RoxygenNote 7.1.2

NeedsCompilation no

Author Hudson Golino [aut, cre] (<https://orcid.org/0000-0002-1601-1447>),
Alexander Christensen [aut] (<https://orcid.org/0000-0002-9798-7037>),
Robert Moulder [ctb] (<https://orcid.org/0000-0001-7504-9560>),
Luis E. Garrido [ctb] (<https://orcid.org/0000-0001-8932-6063>),
Laura Jamison [ctb]

Repository CRAN

Date/Publication 2022-05-10 14:10:02 UTC

R topics documented:

- EGAnet-package .................................................. 3
- boot.ergoInfo .................................................... 5
- boot.wmt .......................................................... 8
- bootEGA ............................................................. 8
- CFA ................................................................. 12
- color_palette_EGA ............................................... 14
- compare.EGA.plots .............................................. 15
- depression ......................................................... 16
- dimensionStability ............................................... 17
- dnn.weights ....................................................... 19
- dynEGA ............................................................. 19
- dynEGA.ind.pop ................................................ 22
- EBICglasso.qgraph ............................................... 24
- EGA ................................................................. 26
- EGA.estimate ..................................................... 30
- EGA.fit .............................................................. 33
- ega.wmt ............................................................. 36
- Embed .............................................................. 36
- entropyFit ......................................................... 37
- ergoInfo ............................................................ 39
- glia ................................................................. 40
- hierEGA ............................................................. 41
- intelligenceBattery .............................................. 45
- invariance ......................................................... 46
- itemStability ...................................................... 47
- LCT ................................................................. 49
- louvain ............................................................. 51
- mctest.ergoInfo .................................................. 52
- methods.section .................................................. 56
Description

Implements the Exploratory Graph Analysis (EGA; Golino & Epskamp, 2017; Golino, Shi et al., 2020) framework for dimensionality and psychometric assessment. EGA is part of a new area called network psychometrics that uses undirected network models for the assessment of psychometric properties. EGA estimates the number of dimensions (or factors) using graphical lasso EBICglasso or Triangulated Maximally Filtered Graph (TMFG) and a weighted network community detection algorithm (Christensen, Garrido, Golino, under review A). A bootstrap method for verifying the stability of the dimensions and items in those dimensions is available (bootEGA; Christensen & Golino, 2021a). The fit of the structure suggested by EGA can be verified using Entropy Fit Indices (entropyFit, tefi; Golino, Moulder et al., 2020). A novel approach called Unique Variable Analysis (UVA) can be used to identify and reduce redundant variables in multivariate data (Christensen, Garrido, & Golino, under review B). Network loadings (net.loads), which are roughly equivalent to factor loadings when the data generating model is a factor model, are available (Christensen & Golino, 2021b, 2021c). Network scores (net.scores) can also be computed using the network loadings. Finally, dynamic EGA (dynEGA) will estimate dimensions from time series data for individual, group, and sample levels (Golino, Christensen et al., 2021).

Author(s)

Hudson Golino <hfg9s@virginia.edu> and Alexander P. Christensen <alexpaulchristensen@gmail.com>
References

# Related functions: EGA

# Related functions: UVA

# Related functions: bootEGA, dimensionStability, and itemStability

# Related functions: LCT

# Related functions: LCT and net.loads

# Related functions: bootEGA, dimensionStability, EGA, itemStability, and UVA

# Related functions: dynEGA and simDFM

# Related functions: EGA

# Related functions: EGA

# Related functions: entropyFit, tefi, and vn.entropy

# Related functions: EGA

# Related functions: EGA.fit and tefi
boot.ergoInfo


# Related functions: EGA.fit and tefi

---

### boot.ergoInfo

**Bootstrap Test for the Ergodicity Information Index**

#### Description

Computes a parametric Bootstrap Test for the Ergodicity Information Index, comparing the empirical Ergodicity Information index to values obtained in data generated using N parametric bootstraps of the correlation matrix estimated using the dynEGA function, for the population structure. The p-values in the bootstrap test can be calculated as \((\text{sum}(\text{EII}>=\text{boot.EII})+1)/(\text{iter}+1)\) and as \((\text{sum}(\text{EII}<\text{boot.EII})+1)/(\text{iter}+1)\), where EII is the empirical Ergodicity Information Index, boot.EII is the values of the Ergodicity Information Index obtained in the bootstraped samples, and iter is the number of random samples generated in the simulation. The two-sided p-value is computed as two times the lowest p-value. In the bootstrap Test for the Ergodicity Information Index, the null hypothesis is that the empirical value of EII is equal to the values of EII obtained in multiple individuals with the same structure as the population structure estimated via dynEGA. Small values of \(p\) indicate that is very unlikely to obtain an EII as large as the one obtained in the empirical sample if the null hypothesis is true (i.e. all individuals have the same structure as the population structure), thus there is convincing evidence that the empirical Ergodicity Information Index is different than it could be expected if all individuals had a similar latent structure.

#### Usage

```r
boot.ergoInfo(
  dynEGA.pop, 
  iter, 
  EII, 
  use, 
  embed, 
  tau, 
  delta, 
  derivatives, 
  model, 
  model.args = list(), 
  algorithm = c("walktrap", "louvain"), 
  algorithm.args = list(), 
  corr, 
  ncores, 
  ... 
)
```

#### Arguments

- `dynEGA.pop`: A dynEGA or a dynEGA.pop.ind object.
**iter**  Numeric integer. Number of random samples to generate in the Monte-Carlo simulation. At least 500 is recommended

**EII**  Numeric. Empirical Ergodicity Information Index obtained via the `ergoInfo` function.

**use**  Character. A string indicating what network element will be used to compute the algorithm complexity in the `ergoInfo` function, the list of edges or the weights of the network. Defaults to use = "edge.list". Current options are:

- `edge.list` Calculates the algorithm complexity using the list of edges.
- `weights` Calculates the algorithm complexity using the weights of the network.

**embed**  Integer. Number of embedded dimensions (the number of observations to be used in the `Embed` function). For example, an "embed = 5" will use five consecutive observations to estimate a single derivative. Default is "embed = 5".

**tau**  Integer. Number of observations to offset successive embeddings in the `Embed` function. A tau of one uses adjacent observations. Default is "tau = 1".

**delta**  Integer. The time between successive observations in the time series. Default is "delta = 1".

**derivatives**  Integer. The order of the derivative to be used in the EGA procedure. Default to 1.

**model**  Character. A string indicating the method to use. Defaults to `glasso`. Current options are:

- `glasso` Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter. This is the default method
- `TMFG` Estimates a Triangulated Maximally Filtered Graph

**model.args**  List. A list of additional arguments for `EBICglasso.qgraph` or `TMFG`

**algorithm**  A string indicating the algorithm to use or a function from `igraph`

**algorithm.args**  List. A list of additional arguments for `cluster_walktrap`, `cluster_louvain`, or some other community detection algorithm function (see examples)

**corr**  Type of correlation matrix to compute. The default uses `cor_auto`. Current options are:

- `cor_auto` Computes the correlation matrix using the `cor_auto` function from `qgraph`
- `pearson` Computes Pearson’s correlation coefficient using the pairwise complete observations via the `cor` function
- `spearman` Computes Spearman’s correlation coefficient using the pairwise complete observations via the `cor` function.
`boot.ergoInfo`

**ncores**

Numeric. Number of cores to use in computing results. Defaults to `parallel::detectCores() / 2` or half of your computer’s processing power. Set to 1 to not use parallel computing. Recommended to use maximum number of cores minus one.

If you’re unsure how many cores your computer has, then use the following code: `parallel::detectCores()`

Additional arguments. Used for deprecated arguments from previous versions of EGA.

**Value**

Returns a list containing:

- `boot.ergoInfo`: The values of the Ergodicity Information Index obtained in the Monte-Carlo Simulation.
- `p.value.twosided`: The p-value of the Monte-Carlo test for the Ergodicity Information Index. The null hypothesis is that the empirical Ergodicity Information index is equal to the expected value of the EII if the all individuals had similar latent structures.
- `effect`: Indicates whether the empirical EII is greater or less than the Monte-Carlo obtained EII.
- `plot.dist`: Histogram of the bootstrapped ergodicity information index.

**Author(s)**

Hudson Golino <hfg9s at virginia.edu>

**Examples**

```r
## Not run:
\donttest{
dyn1 <- dynEGA.ind.pop(data = sim.dynEGA[, -c(22)], n.embed = 5, tau = 1,
                       delta = 1, id = 21, use.derivatives = 1,
                       model = "glasso", ncores = 2, corr = "pearson")
eii1 <- ergoInfo(data = dyn1)$EII
testing.ergoinfo <- boot.ergoInfo(dynEGA.pop = dyn1, iter = 10, EII = eii1,
                                 embed = 5, tau = 1, delta = 1, derivatives = 1,
                                 model = "glasso", ncores = 2, corr = "pearson")
}
## End(Not run)
```
Description

bootEGA results using the "glasso" model and 500 iterations of the Wiener Matrizen-Test 2 (WMT-2)

Examples
data("boot.wmt")
data("boot.wmt")

bootEGA

Dimension Stability Analysis of EGA

Description

bootEGA Estimates the number of dimensions of n bootstraps using the empirical (partial) correlation matrix (parametric) or resampling from the empirical dataset (non-parametric). It also estimates a typical median network structure, which is formed by the median or mean pairwise (partial) correlations over the n bootstraps.
Usage

```r
bootEGA(
  data,
  n = NULL,
  uni.method = c("expand", "LE"),
  iter,
  type = c("parametric", "resampling"),
  seed = 1234,
  corr = c("cor_auto", "pearson", "spearman"),
  model = c("glasso", "TMFG"),
  model.args = list(),
  algorithm = c("walktrap", "louvain"),
  algorithm.args = list(),
  typicalStructure = TRUE,
  plot.typicalStructure = TRUE,
  plot.type = c("GGally", "qgraph"),
  plot.args = list(),
  ncores,
  ...
)
```

Arguments

data       Matrix or data frame. Includes the variables to be used in the bootEGA analysis
n          Integer. Sample size if data provided is a correlation matrix
uni.method Character. What unidimensionality method should be used? Defaults to "LE". Current options are:
• expand Expands the correlation matrix with four variables correlated .50. If number of dimension returns 2 or less in check, then the data are unidimensional; otherwise, regular EGA with no matrix expansion is used. This is the method used in the Golino et al. (2020) Psychological Methods simulation.
• LE Applies the leading eigenvalue algorithm (`cluster_leading_eigen`) on the empirical correlation matrix. If the number of dimensions is 1, then the leading eigenvalue solution is used; otherwise, regular EGA is used. This is the final method used in the Christensen, Garrido, and Golino (2021) simulation.
iter       Numeric integer. Number of replica samples to generate from the bootstrap analysis. At least 500 is recommended
type       Character. A string indicating the type of bootstrap to use. Current options are:
• "parametric" Generates `n` new datasets (multivariate normal random distributions) based on the original dataset, via the `mvnorm` function
• "resampling" Generates random subsamples of the original data
seed       Numeric. Seed to reproduce results. Defaults to 1234. For random results, set to NULL
corr

Type of correlation matrix to compute. The default uses cor_auto. Current options are:

• cor_auto Computes the correlation matrix using the cor_auto function from qgraph.
• pearson Computes Pearson’s correlation coefficient using the pairwise complete observations via the cor function.
• spearman Computes Spearman’s correlation coefficient using the pairwise complete observations via the cor function.

model

Character. A string indicating the method to use. Current options are:

• glasso Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter. This is the default method
• TMFG Estimates a Triangulated Maximally Filtered Graph

model.args

List. A list of additional arguments for EBICglasso.qgraph or TMFG

algorithm

A string indicating the algorithm to use or a function from igraph Current options are:

• walktrap Computes the Walktrap algorithm using cluster_walktrap
• louvain Computes the Walktrap algorithm using cluster_louvain

algorithm.args

List. A list of additional arguments for cluster_walktrap, cluster_louvain, or some other community detection algorithm function (see examples)

typicalStructure

Boolean. If TRUE, returns the typical network of partial correlations (estimated via graphical lasso or via TMFG) and estimates its dimensions. The "typical network" is the median of all pairwise correlations over the n bootstraps. Defaults to TRUE

plot.typicalStructure

Boolean. If TRUE, returns a plot of the typical network (partial correlations), which is the median of all pairwise correlations over the n bootstraps, and its estimated dimensions. Defaults to TRUE

plot.type

Character. Plot system to use. Current options are qgraph and GGally. Defaults to "GGally".

plot.args

List. A list of additional arguments for the network plot. For plot.type = "qgraph":

• vsize Size of the nodes. Defaults to 6.
For plot.type = "GGally" (see ggnet2 for full list of arguments):

• vsize Size of the nodes. Defaults to 6.
• label.size Size of the labels. Defaults to 5.
• alpha The level of transparency of the nodes, which might be a single value or a vector of values. Defaults to 0.7.
• edge.alpha The level of transparency of the edges, which might be a single value or a vector of values. Defaults to 0.4.
bootEGA

- legend.names A vector with names for each dimension
- color.palette The color palette for the nodes. For custom colors, enter HEX codes for each dimension in a vector. See color_palette_EGA for more details and examples

tncores Numeric. Number of cores to use in computing results. Defaults to parallel::detectCores() / 2 or half of your computer’s processing power. Set to 1 to not use parallel computing

If you’re unsure how many cores your computer has, then use the following code: parallel::detectCores()

... Additional arguments. Used for deprecated arguments from previous versions of EGA

Value

Returns a list containing:

iter Number of replica samples in bootstrap
boot.ndim Number of dimensions identified in each replica sample
boot.wc Item allocation for each replica sample
bootGraphs Networks of each replica sample
summary.table Summary table containing number of replica samples, median, standard deviation, standard error, 95% confidence intervals, and quantiles (lower = 2.5% and upper = 97.5%)
frequency Proportion of times the number of dimensions was identified (e.g., .85 of 1,000 = 850 times that specific number of dimensions was found)
EGA Output of the original EGA results
typicalGraph A list containing:
  - graph Network matrix of the median network structure
  - typical.dim.variables An ordered matrix of item allocation
  - wc Item allocation of the median network

Author(s)

Hudson Golino <hfg9s at virginia.edu> and Alexander P. Christensen <alexpaulchristensen@gmail.com>

References

# Original implementation of bootEGA

# Structural consistency (see dimensionStability)
See Also

EGA to estimate the number of dimensions of an instrument using EGA and CFA to verify the fit of the structure suggested by EGA using confirmatory factor analysis.

Examples

# Load data
wmt <- wmt2[,7:24]

# bootEGA glasso example
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
boot.wmt <- bootEGA(data = wmt, iter = 500, plot.type = "qgraph",
type = "parametric", ncores = 2)

# bootEGA TMFG example
boot.wmt <- bootEGA(data = wmt, iter = 500, model = "TMFG",
plot.type = "qgraph", type = "parametric", ncores = 2, seed = 1234)

# bootEGA Louvain example
boot.wmt <- bootEGA(data = wmt, iter = 500, algorithm = "louvain",
plot.type = "qgraph", type = "parametric", ncores = 2, seed = 1234)

# bootEGA Spinglass example
boot.wmt <- bootEGA(data = wmt, iter = 500, model = "TMFG",
algorithm = igraph::cluster_spinglass, type = "parametric", ncores = 2)

---

CFA

CFA Fit of EGA Structure

Description

Verifies the fit of the structure suggested by EGA using confirmatory factor analysis

Usage

CFA(ega.obj, data, estimator, plot.CFA = TRUE, layout = "spring", ...)

Arguments

ega.obj An EGA object
data A dataframe with the variables to be used in the analysis
estimator The estimator used in the confirmatory factor analysis. 'WLSMV' is the estimator of choice for ordinal variables. 'ML' or 'WLS' for interval variables. See lavOptions for more details
plot.CFA Logical. Should the CFA structure with its standardized loadings be plot? Defaults to `TRUE`

layout Layout of plot (see `semPaths`). Defaults to "spring"

Arguments passed to `cfa`

Value

Returns a list containing:

- `fit` Output from `cfa`
- `summary` Summary output from `lavaan-class`
- `fit.measures` Fit measures: chi-squared, degrees of freedom, p-value, CFI, RMSEA, GFI, and NFI. Additional fit measures can be applied using the `fitMeasures` function (see examples)

Author(s)

Hudson F. Golino <hfg9s at virginia.edu>

References


See Also

- `EGA` to estimate the number of dimensions of an instrument using EGA and `bootEGA` to investigate the stability of EGA's estimation via bootstrap.

Examples

```r
# Load data
wmt <- wmt2[,7:24]

# Estimate EGA
g1 <- EGA(data = wmt, plot.EGA = FALSE)

g2 <- CFA(ega.obj = g1, estimator = 'WLSMV', plot.CFA = TRUE, data = wmt)

# Additional fit measures
lavaan::fitMeasures(g2$fit, fit.measures = "all")
```

# Load data
```r
intel <- intelligenceBattery[,8:66]

# Estimate EGA
ega.intel <- EGA(data = intel, plot.EGA = FALSE)

# Fit CFA model to EGA results
cfa.intel <- CFA(ega.obj = ega.intel, estimator = 'WLSMV', plot.CFA = TRUE,
data = intel)
```

---

**color_palette_EGA**

**EGA Color Palettes**

**Description**

Color palettes for plotting ggnet2 EGA network plots

**Usage**

```
color_palette_EGA(name, wc, sorted = FALSE)
```

**Arguments**

- `name` Character. Name of color scheme (see RColorBrewer). Defaults to "polychrome". EGA palettes:
  - "polychrome" Default 20 color palette
  - "grayscale" "grayscale", "greyscale", or "colorblind" will produce plots suitable for publication purposes
  - "blue.ridge1" Palette inspired by the Blue Ridge Mountains
  - "blue.ridge2" Second palette inspired by the Blue Ridge Mountains
  - "rainbow" Rainbow colors. Default for qgraph
  - "rio" Palette inspired by Rio de Janiero, Brazil
  - "itacare" Palette inspired by Itacare, Brazil

  For custom colors, enter HEX codes for each dimension in a vector

- `wc` Vector. A vector representing the community (dimension) membership of each node in the network. NA values mean that the node was disconnected from the network

- `sorted` Boolean. Should colors be sorted by wc? Defaults to TRUE

**Value**

Vector of colors for community memberships
Author(s)
Hudson Golino <hfg9s at virginia.edu>, Alexander P. Christensen <alexpaulchristensen at gmail.com>

Examples

# Default
color_palette_EGA(name = "polychrome", wc = ega.wmt$wc)

# Blue Ridge Moutains 1
color_palette_EGA(name = "blue.ridge1", wc = ega.wmt$wc)

# Custom
color_palette_EGA(name = "#7FD1B9", wc = ega.wmt$wc)

Description
Organizes EGA plots for comparison. Ensures that nodes are placed in the same layout to maximize comparison. Community memberships are also homogenized across EGA outputs to enhance interpretation.

Usage

compare.EGA.plots(..., input_list = NULL, base_plot = 1, labels, rows, columns, plot.type = c("GGally", "qgraph"), plot.args = list())

Arguments

... EGAnet objects
input_list List. Bypasses ... argument in favor of using a list as an input
base_plot Numeric. Plot to be used as the base for the configuration of the networks. Uses the number of the order in which the plots are input. Defaults to 1 or the first plot
labels Character vector. Labels for each EGAnet object
rows Numeric. Number of rows to spread plots across
columns Numeric. Number of columns to spread plots down

plot.type Character. Plot system to use. Current options are qgraph and GGally. Defaults to "GGally"

plot.args List. A list of additional arguments for the network plot. For plot.type = "qgraph":
  • vsiz e Size of the nodes. Defaults to 6.
For plot.type = "GGally" (see ggnet2 for full list of arguments):
  • vsiz e Size of the nodes. Defaults to 6.
  • label.size Size of the labels. Defaults to 5.
  • alpha The level of transparency of the nodes, which might be a single value or a vector of values. Defaults to 0.7.
  • edge.alpha The level of transparency of the edges, which might be a single value or a vector of values. Defaults to 0.4.
  • legend.names A vector with names for each dimension
  • color.palette The color palette for the nodes. For custom colors, enter HEX codes for each dimension in a vector. See color_palette_EGA for more details and examples

Value

Visual comparison of EGAnet objects

Author(s)

Alexander Christensen <alexpaulchristensen@gmail.com>

Examples

# obtain SAPA items
items <- psychTools::spi[,c(11:20)]

---

**depression**

*Depression Data*

Description

A response matrix (n = 574) of the Beck Depression Inventory, Beck Anxiety Inventory and the Athens Insomnia Scale.

Usage

data(depression)

data(depression)
**dimensionStability**

**Format**
A 574x78 response matrix
A 574x78 response matrix

**Examples**
```
data("depression")
data("depression")
```

---

**dimensionStability**  
*Dimension Stability Statistics from bootEGA*

**Description**
Based on the bootEGA results, this function computes the stability of dimensions. This is computed by assessing the proportion of times the original dimension is exactly replicated in across bootstrap samples.

**Usage**
```
dimensionStability(bootega.obj, ...)
```

**Arguments**
- `bootega.obj`  
  A bootEGA object

- `...`  
  Additional arguments. Used for deprecated arguments from previous versions of dimStability.

**Value**
Returns a list containing:
- `dimension.stability`  
  A list containing:
  - `structural.consistency` The proportion of times that each empirical EGA dimension exactly replicates across the bootEGA samples
  - `average.item.stability` The average item stability in each empirical EGA dimension

- `item.stability`  
  Results from itemStability

**Author(s)**
Hudson Golino <hfg9s at virginia.edu> and Alexander P. Christensen <alexpaulchristensen@gmail.com>
References


See Also

EGA to estimate the number of dimensions of an instrument using EGA and CFA to verify the fit of the structure suggested by EGA using confirmatory factor analysis.

Examples

```r
# Load data
wmt <- wmt2[,7:24]

## Not run: # Estimate EGA network
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
ga.wmt <- EGA(data = wmt, model = "glasso", plot.type = "qgraph")

# Estimate bootstrap EGA
boot.wmt <- bootEGA(data = wmt, iter = 500, typicalStructure = TRUE, plot.typicalStructure = TRUE, model = "glasso", plot.type = "qgraph", type = "parametric", ncores = 2)

## End(Not run)

# Estimate stability statistics
res <- dimensionStability(boot.wmt)
res$dimension.stability

# Changing plot features (ggplot2)
## Changing colors (ignore warnings)
### qgraph Defaults
res$item.stability$plot +
  ggplot2::scale_color_manual(values = rainbow(length(res$dimension.stability$structural.consistency)))

### Pastel
res$item.stability$plot +
  ggplot2::scale_color_brewer(palette = "Pastel1")

## Changing Legend (ignore warnings)
res$item.stability$plot +
  ggplot2::scale_color_discrete(labels = "Intelligence")
```
**dnn.weights**

### Loadings Comparison Test Deep Learning Neural Network Weights

#### Description

A list of weights from four different neural network models: random vs. non-random model ($r_{nr}$), low correlation factor vs. network model ($lf_{n}$), high correlation with variables less than or equal to factors vs. network model ($hlf_{n}$), and high correlation with variables greater than factors vs. network model ($hgf_{n}$).

#### Usage

```r
data(dnn.weights)
```

#### Format

A list of with a length of 4

#### Examples

```r
data("dnn.weights")
data("dnn.weights")
```

---

**dynEGA**

### Dynamic Exploratory Graph Analysis

#### Description

Estimates dynamic factors in multivariate time series (i.e. longitudinal data, panel data, intensive longitudinal data) at multiple time scales, in different levels of analysis: individuals (intraindividual structure), groups or population (structure of the population). Exploratory graph analysis is applied in the derivatives estimated using generalized local linear approximation (glla). Instead of estimating factors by modeling how variables are covarying, as in traditional EGA, dynEGA is a dynamic model that estimates the factor structure by modeling how variables are changing together. GLLA is a filtering method for estimating derivatives from data that uses time delay embedding and a variant of Savitzky-Golay filtering to accomplish the task.
Usage

dynEGA(
  data,  # A dataframe with the variables to be used in the analysis. The dataframe should
  n.embed,  # be in a long format (i.e. observations for the same individual (for example, individual 1)
  tau = 1,  # are placed in order, from time 1 to time t, followed by the observations from
  delta = 1,  # individual 2, also ordered from time 1 to time t.)
  level = c("individual", "group", "population"),  # Character. A string indicating the level of analysis. If the interest is in modeling
  id = NULL,  # the intraindividual structure only (one dimensionality structure per individual),
  group = NULL,  # then level should be set to "individual". If the interest is in the structure of a
  use.derivatives = 1,  # group of individuals, then level should be set to "group". Finally, if the interest
  model = c("glasso", "TMFG"),  # is in the population structure, then level should be set to "population".
  model.args = list(),  # Current options are:
  algorithm = c("walktrap", "louvain"),  # • individual Estimates the dynamic factors per individual. This should be
  algorithm.args = list(),  # the preferred method is one is interested in in the factor structure of individuals. An additional parameter ("id") needs to be provided identifying each
  corr = c("cor_auto", "pearson", "spearman"),  # individual.
  ncores,  # • group Estimates the dynamic factors for each group. An additional parameter ("group") needs to be provided identifying the group membership.
  ...  # • population Estimates the dynamic factors of the population
)

Arguments

data  # A dataframe with the variables to be used in the analysis. The dataframe should
n.embed  # be in a long format (i.e. observations for the same individual (for example, individual 1)
tau  # are placed in order, from time 1 to time t, followed by the observations from
delta  # individual 2, also ordered from time 1 to time t.)
level  # Character. A string indicating the level of analysis. If the interest is in modeling
id  # the intraindividual structure only (one dimensionality structure per individual),
group  # then level should be set to "individual". If the interest is in the structure of a
use.derivatives  # group of individuals, then level should be set to "group". Finally, if the interest
model  # is in the population structure, then level should be set to "population".
model.args  # Current options are:
algorithm  # • individual Estimates the dynamic factors per individual. This should be
algorithm.args  # the preferred method is one is interested in in the factor structure of individuals. An additional parameter ("id") needs to be provided identifying each
corr  # • group Estimates the dynamic factors for each group. An additional parameter ("group") needs to be provided identifying the group membership.
ncores  # • population Estimates the dynamic factors of the population
...  # • individual Estimates the dynamic factors per individual. This should be
  # the preferred method is one is interested in in the factor structure of individuals. An additional parameter ("id") needs to be provided identifying each individual.
  # • group Estimates the dynamic factors for each group. An additional parameter ("group") needs to be provided identifying the group membership.
  # • population Estimates the dynamic factors of the population
id  # Numeric. Number of the column identifying each individual.
dynEGA

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
<td>Numeric or character. Number of the column identifying group membership. Must be specified only if level = &quot;group&quot;.</td>
</tr>
<tr>
<td>use.derivatives</td>
<td>Integer. The order of the derivative to be used in the EGA procedure. Default to 1.</td>
</tr>
<tr>
<td>model</td>
<td>Character. A string indicating the method to use. Current options are: - glasso Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter. This is the default method - TMFG Estimates a Triangulated Maximally Filtered Graph</td>
</tr>
<tr>
<td>model.args</td>
<td>List. A list of additional arguments for EBICglasso.qgraph or TMFG algorithm.</td>
</tr>
<tr>
<td>algorithm</td>
<td>A string indicating the algorithm to use or a function from igraph algorithm.args List. A list of additional arguments for cluster_walktrap, cluster_louvain, or some other community detection algorithm function (see examples)</td>
</tr>
<tr>
<td>corr</td>
<td>Type of correlation matrix to compute. The default uses &quot;pearson&quot;. Current options are: - cor_auto Computes the correlation matrix using the cor_auto function from qgraph. - pearson Computes Pearson’s correlation coefficient using the pairwise complete observations via the cor function. - spearman Computes Spearman’s correlation coefficient using the pairwise complete observations via the cor function.</td>
</tr>
<tr>
<td>ncores</td>
<td>Numeric. Number of cores to use in computing results. Defaults to parallel::detectCores() / 2 or half of your computer’s processing power. Set to 1 to not use parallel computing. Recommended to use maximum number of cores minus one</td>
</tr>
<tr>
<td>...</td>
<td>Additional arguments. Used for deprecated arguments from previous versions of EGA</td>
</tr>
</tbody>
</table>

**Author(s)**

Hudson Golino <hfg9s at virginia.edu>

**References**


**Examples**

```r
# Population structure:
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
dyn.random <- dynEGA(data = sim.dynEGA, n.embed = 5, tau = 1,
delta = 1, id = 21, group = 22, use.derivatives = 1,
level = "population", model = "glasso", ncores = 2)

plot(dyn.random, plot.type = "qgraph")

# Group structure:
dyn.group <- dynEGA(data = sim.dynEGA, n.embed = 5, tau = 1,
delta = 1, id = 21, group = 22, use.derivatives = 1,
level = "group", model = "glasso", ncores = 2)

plot(dyn.group, ncol = 2, nrow = 1, plot.type = "qgraph")

# Intraindividual structure (commented out for CRAN tests):
# dyn.individual <- dynEGA(data = sim.dynEGA, n.embed = 5, tau = 1,
# delta = 1, id = 21, group = 22, use.derivatives = 1,
# level = "individual", model = "glasso", ncores = 2)
```

---

**Description**

Dynamic EGA used in the mctest.ergoInfo function. DynEGA estimates dynamic factors in multivariate time series (i.e. longitudinal data, panel data, intensive longitudinal data) at multiple time scales, in different levels of analysis: individuals (intraindividual structure) and population (structure of the population). Exploratory graph analysis is applied in the derivatives estimated using generalized local linear approximation (*glla*). Instead of estimating factors by modeling how variables are covarying, as in traditional EGA, dynEGA is a dynamic model that estimates the factor structure by modeling how variables are changing together. *Glla* is a filtering method for estimating derivatives from data that uses time delay embedding and a variant of Savitzky-Golay filtering to accomplish the task.
Usage

dynEGA.ind.pop(
  data,
  n.embed,
  tau = 1,
  delta = 1,
  id = NULL,
  use.derivatives = 1,
  model = c("glasso", "TMFG"),
  model.args = list(),
  algorithm = c("walktrap", "louvain"),
  algorithm.args = list(),
  corr = c("cor_auto", "pearson", "spearman"),
  ncores,
  ...
)

Arguments

data A dataframe with the variables to be used in the analysis. The dataframe should
be in a long format (i.e. observations for the same individual (for example, indi-
vidual 1) are placed in order, from time 1 to time t, followed by the observations
from individual 2, also ordered from time 1 to time t.)
n.embed Integer. Number of embedded dimensions (the number of observations to be
used in the Embed function). For example, an "n.embed = 5" will use five con-
secutive observations to estimate a single derivative.
tau Integer. Number of observations to offset successive embeddings in the Embed
function. A tau of one uses adjacent observations. Default is "tau = 1".
delta Integer. The time between successive observations in the time series. Default is
"delta = 1".
id Numeric. Number of the column identifying each individual.
use.derivatives Integer. The order of the derivative to be used in the EGA procedure. Default to
1.
model Character. A string indicating the method to use. Defaults to glasso. Current
options are:
  • glasso Estimates the Gaussian graphical model using graphical LASSO
    with extended Bayesian information criterion to select optimal regulariza-
    tion parameter. This is the default method
  • TMFG Estimates a Triangulated Maximally Filtered Graph
model.args List. A list of additional arguments for EBICglasso.qgraph or TMFG
algorithm A string indicating the algorithm to use or a function from igraph
Current options are:
  • walktrap Computes the Walktrap algorithm using cluster_walktrap
  • louvain Computes the Walktrap algorithm using cluster_louvain
algorithm.args List. A list of additional arguments for `cluster_walktrap`, `cluster_louvain`, or some other community detection algorithm function (see examples)

corr Type of correlation matrix to compute. The default uses `cor_auto`. Current options are:

- `cor_auto` Computes the correlation matrix using the `cor_auto` function from `qgraph`.
- `pearson` Computes Pearson’s correlation coefficient using the pairwise complete observations via the `cor` function.
- `spearman` Computes Spearman’s correlation coefficient using the pairwise complete observations via the `cor` function.

ncores Numeric. Number of cores to use in computing results. Defaults to `parallel::detectCores()` / 2 or half of your computer’s processing power. Set to 1 to not use parallel computing. Recommended to use maximum number of cores minus one

If you’re unsure how many cores your computer has, then use the following code: `parallel::detectCores()`

Additional arguments. Used for deprecated arguments from previous versions of EGA

Author(s)

Hudson Golino <hfg9s at virginia.edu>

Examples

```r
## Not run:
\donttest{

dyn.ega1 <- dynEGA.ind.pop(data = sim.dynEGA, n.embed = 5, tau = 1, 
delta = 1, id = 21, use.derivatives = 1, model = "glasso", ncores = 2, 
cor = "pearson")
}
## End(Not run)
```

Description

This function uses the `glasso` package (Friedman, Hastie and Tibshirani, 2011) to compute a sparse gaussian graphical model with the graphical lasso (Friedman, Hastie & Tibshirani, 2008). The tuning parameter is chosen using the Extended Bayesian Information criterium (EBIC) described by Foygel & Drton (2010).
Usage

EBICglasso.qgraph(
  data,
  n = NULL,
  gamma = 0.5,
  penalize.diagonal = FALSE,
  nlambda = 100,
  lambda.min.ratio = 0.01,
  returnAllResults = FALSE,
  penalizeMatrix,
  countDiagonal = FALSE,
  refit = FALSE,
...)

Arguments

data  Data matrix
n      Number of participants
gamma EBIC tuning parameter. 0.5 is generally a good choice. Setting to zero will cause regular BIC to be used.
penalize.diagonal  Should the diagonal be penalized?
nlambda Number of lambda values to test.
lambda.min.ratio Ratio of lowest lambda value compared to maximal lambda
returnAllResults  If TRUE this function does not return a network but the results of the entire glasso path.
penalizeMatrix  Optional logical matrix to indicate which elements are penalized
countDiagonal  Should diagonal be counted in EBIC computation? Defaults to FALSE. Set to TRUE to mimic qgraph < 1.3 behavior (not recommended!).
refit  Logical, should the optimal graph be refitted without LASSO regularization? Defaults to FALSE.
...  Arguments sent to glasso

Details

The glasso is run for 100 values of the tuning parameter logarithmically spaced between the maximum value of the tuning parameter at which all edges are zero, lambda_max, and lambda_max/100. For each of these graphs the EBIC is computed and the graph with the best EBIC is selected. The partial correlation matrix is computed using wi2net and returned.

Value

A partial correlation matrix
Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

References


Examples

```r
### Using wmt2 dataset from EGAnet ###
data(wmt2)

# Compute correlations:
CorMat <- qgraph::cor_auto(wmt2[,7:24])

# Compute graph with tuning = 0 (BIC):
BICgraph <- EBICglasso.qgraph(CorMat, n = nrow(wmt2), gamma = 0)

# Compute graph with tuning = 0.5 (EBIC)
EBICgraph <- EBICglasso.qgraph(CorMat, n = nrow(wmt2), gamma = 0.5)
```

EGA

*Applies the Exploratory Graph Analysis technique*

Description

Estimates the number of dimensions of a given dataset or correlation matrix using the graphical lasso (*EBICglasso.qgraph*) or the Triangulated Maximally Filtered Graph (*TMFG*) network estimation methods.
EGA

Usage

EGA(
  data,
  n = NULL,
  uni.method = c("expand", "LE"),
  corr = c("cor_auto", "pearson", "spearman"),
  model = c("glasso", "TMFG"),
  model.args = list(),
  algorithm = c("walktrap", "louvain"),
  algorithm.args = list(),
  plot.EGA = TRUE,
  plot.type = c("GGally", "qgraph"),
  plot.args = list(),
  verbose = TRUE,
  ...
)

Arguments

data  Matrix or data frame. Variables (down columns) or correlation matrix. If the input is a correlation matrix, then argument n (number of cases) is **required**
n  Integer. Sample size if data provided is a correlation matrix
uni.method  Character. What unidimensionality method should be used? Defaults to "LE". Current options are:
  - expand Expands the correlation matrix with four variables correlated .50. If number of dimension returns 2 or less in check, then the data are unidimensional; otherwise, regular EGA with no matrix expansion is used. This is the method used in the Golino et al. (2020) *Psychological Methods* simulation.
  - LE Applies the leading eigenvalue algorithm (*cluster_leading_eigen*) on the empirical correlation matrix. If the number of dimensions is 1, then the leading eigenvalue solution is used; otherwise, regular EGA is used. This is the final method used in the Christensen, Garrido, and Golino (2021) simulation.
corr  Type of correlation matrix to compute. The default uses *cor_auto*. Current options are:
  - cor_auto Computes the correlation matrix using the *cor_auto* function from *qgraph*.
  - pearson Computes Pearson’s correlation coefficient using the pairwise complete observations via the *cor* function.
  - spearman Computes Spearman’s correlation coefficient using the pairwise complete observations via the *cor* function.
model  Character. A string indicating the method to use. Defaults to "glasso". Current options are:
EGA

- glasso Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter
- TMFG Estimates a Triangulated Maximally Filtered Graph

model.args List. A list of additional arguments for EBICglasso.qgraph or TMFG
algorithm A string indicating the algorithm to use or a function from igraph. Defaults to "walktrap". Current options are:
  - walktrap Computes the Walktrap algorithm using cluster_walktrap
  - louvain Computes the Louvain algorithm using cluster_louvain

algorithm.args List. A list of additional arguments for cluster_walktrap, cluster_louvain, or some other community detection algorithm function (see examples)
plot.EGA Boolean. If TRUE, returns a plot of the network and its estimated dimensions. Defaults to TRUE
plot.type Character. Plot system to use. Current options are qgraph and GGally. Defaults to "GGally"
plot.args List. A list of additional arguments for the network plot. For plot.type = "qgraph":
  - vsize Size of the nodes. Defaults to 6.
For plot.type = "GGally" (see ggnet2 for full list of arguments):
  - vsize Size of the nodes. Defaults to 6.
  - label.size Size of the labels. Defaults to 5.
  - alpha The level of transparency of the nodes, which might be a single value or a vector of values. Defaults to 0.7.
  - edge.alpha The level of transparency of the edges, which might be a single value or a vector of values. Defaults to 0.4.
  - legend.names A vector with names for each dimension
  - color.palette The color palette for the nodes. For custom colors, enter HEX codes for each dimension in a vector. See color_palette_EGA for more details and examples
verbose Boolean. Should network estimation parameters be printed? Defaults to TRUE. Set to FALSE for no print out

... Additional arguments. Used for deprecated arguments from previous versions of EGA

Details

Two community detection algorithms, Walktrap (Pons & Latapy, 2006) and Louvain (Blondel et al., 2008), are pre-programmed because of their superior performance in simulation studies on psychological data generated from factor models (Christensen & Golino; 2020; Golino et al., 2020). Notably, any community detection algorithm from the igraph can be used to estimate the number of communities (see examples).
Value

Returns a list containing:

- network: A symmetric network estimated using either the EBICglasso.qgraph or TMFG.
- wc: A vector representing the community (dimension) membership of each node in the network. NA values mean that the node was disconnected from the network.
- n.dim: A scalar of how many total dimensions were identified in the network.
- cor.data: The zero-order correlation matrix.

Author(s)

Hudson Golino <hfg9s at virginia.edu>, Alexander P. Christensen <alexpaulchristensen at gmail.com>, Maria Dolores Nieto <acinodam at gmail.com> and Luis E. Garrido <garrido.luiseduardo at gmail.com>

References

# Louvain algorithm

# Compared all igraph community detections algorithms, introduced Louvain algorithm, simulation with continuous and polytomous data
# Also implements the Leading Eigenvalue unidimensional method

# Original simulation and implementation of EGA


# Current implementation of EGA, introduced unidimensional checks, continuous and dichotomous data

# Walktrap algorithm

See Also

- bootEGA to investigate the stability of EGA’s estimation via bootstrap and CFA to verify the fit of the structure suggested by EGA using confirmatory factor analysis.
Examples

# Estimate EGA
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
ega.wmt <- EGA(data = wmt2[,7:24], plot.type = "qgraph")

# Summary statistics
summary(ega.wmt)

# Estimate EGAtmfg
ega.wmt <- EGA(data = wmt2[,7:24], model = "TMFG", plot.type = "qgraph")

# Estimate EGA with Louvain algorithm
ega.wmt <- EGA(data = wmt2[,7:24], algorithm = "louvain", plot.type = "qgraph")

# Estimate EGA with Spinglass algorithm
ega.wmt <- EGA(data = wmt2[,7:24],
algorithm = igraph::cluster_spinglass, plot.type = "qgraph")

# Estimate EGA
ega.intel <- EGA(data = intelligenceBattery[,8:66], model = "glasso", plot.EGA = FALSE)

# Summary statistics
summary(ega.intel)


EGA.estimate

A Sub-routine Function for EGA

Description

Estimates the number of dimensions of a given dataset or correlation matrix using the graphical lasso (EBICglasso.qgraph) or the Triangulated Maximally Filtered Graph (TMFG) network estimation methods.

Usage

EGA.estimate(
  data,
  n = NULL,
  model = c("glasso", "TMFG"),
  model.args = list(),
  algorithm = c("walktrap", "louvain"),
  algorithm.args = list(),
  corr = c("cor_auto", "pearson", "spearman"),
  verbose = TRUE,
Arguments

- **data**: Matrix or data frame. Variables (down columns) or correlation matrix. If the input is a correlation matrix, then argument `n` (number of cases) is required.
- **n**: Integer. Sample size if data provided is a correlation matrix.
- **model**: Character. A string indicating the method to use. Current options are:
  - `glasso`: Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter. This is the default method.
  - `TMFG`: Estimates a Triangulated Maximally Filtered Graph.
- **model.args**: List. A list of additional arguments for `EBICglasso.qgraph` or `TMFG`.
- **algorithm**: A string indicating the algorithm to use or a function from `igraph`. Current options are:
  - `walktrap`: Computes the Walktrap algorithm using `cluster_walktrap`.
  - `louvain`: Computes the Walktrap algorithm using `cluster_louvain`.
- **algorithm.args**: List. A list of additional arguments for `cluster_walktrap`, `cluster_louvain`, or some other community detection algorithm function (see examples).
- **corr**: Type of correlation matrix to compute. The default uses `cor_auto`. Current options are:
  - `cor_auto`: Computes the correlation matrix using the `cor_auto` function from `qgraph`.
  - `pearson`: Computes Pearson’s correlation coefficient using the pairwise complete observations via the `cor` function.
  - `spearman`: Computes Spearman’s correlation coefficient using the pairwise complete observations via the `cor` function.
- **verbose**: Boolean. Should network estimation parameters be printed? Defaults to TRUE. Set to FALSE for no print out.

Details

Two community detection algorithms, Walktrap (Pons & Latapy, 2006) and Louvain (Blondel et al., 2008), are pre-programmed because of their superior performance in simulation studies on psychological data generated from factor models (Christensen & Golino; 2020; Golino et al., 2020). Notably, any community detection algorithm from the `igraph` can be used to estimate the number of communities (see examples).
EGA.estimate

Value

Returns a list containing:

- **estimated.network**
  A symmetric network estimated using either the EBICglasso.qgraph or TMFG
- **wc**
  A vector representing the community (dimension) membership of each node in the network. NA values mean that the node was disconnected from the network.
- **n.dim**
  A scalar of how many total dimensions were identified in the network.
- **cor.data**
  The zero-order correlation matrix

Author(s)

Alexander P. Christensen <alexpaulchristensen at gmail.com> and Hudson Golino <hfg9s at virginia.edu>

References

# Louvain algorithm

# Compared all igraph community detections algorithms, introduced Louvain algorithm, simulation with continuous and polytomous data

# Original simulation and implementation of EGA


# Current implementation of EGA, introduced unidimensional checks, continuous and dichotomous data

# Walktrap algorithm

See Also

*bootEGA* to investigate the stability of EGA’s estimation via bootstrap and *CFA* to verify the fit of the structure suggested by EGA using confirmatory factor analysis.
Examples

# Estimate EGA
ega.wmt <- EGA.estimate(data = wmt2[,7:24], model = "glasso")

# Estimate EGAtmfg
ega.wmt <- EGA.estimate(data = wmt2[,7:24], model = "TMFG")

# Estimate EGA with Spinglass
ega.wmt <- EGA.estimate(data = wmt2[,7:24], model = "glasso",
algorithm = igraph::cluster_spinglass)

EGA.fit

EGA Optimal Model Fit using the Total Entropy Fit Index (tefi)

Description

Estimates the best fitting model using EGA. The number of steps in the cluster_walktrap detection algorithm is varied and unique community solutions are compared using tefi.

Usage

EGA.fit(
  data,
  n = NULL,
  uni.method = c("expand", "LE"),
  corr = c("cor_auto", "pearson", "spearman"),
  model = c("glasso", "TMFG"),
  algorithm = c("leiden", "walktrap"),
  algorithm.args = list(steps = c(3:8), resolution_parameter = seq(0, 2, 0.001))
)

Arguments

data Matrix or data frame. Dataset or correlation matrix
n Integer. Sample size (if the data provided is a correlation matrix)
uni.method Character. What unidimensionality method should be used? Defaults to "LE". Current options are:
  • expand Expands the correlation matrix with four variables correlated .50. If number of dimension returns 2 or less in check, then the data are unidimensional; otherwise, regular EGA with no matrix expansion is used. This is the method used in the Golino et al. (2020) Psychological Methods simulation.
EGA.fit

- LE Applies the leading eigenvalue algorithm (cluster_leading_eigen) on the empirical correlation matrix. If the number of dimensions is 1, then the leading eigenvalue solution is used; otherwise, regular EGA is used. This is the final method used in the Christensen, Garrido, and Golino (2021) simulation.

corr Type of correlation matrix to compute. The default uses cor_auto. Current options are:
  - cor_auto Computes the correlation matrix using the cor_auto function from qgraph.
  - pearson Computes Pearson’s correlation coefficient using the pairwise complete observations via the cor function.
  - spearman Computes Spearman’s correlation coefficient using the pairwise complete observations via the cor function.

model Character. A string indicating the method to use. Defaults to "glasso". Current options are:
  - "glasso" Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter. See EBICglasso.qgraph
  - "TMFG" Estimates a Triangulated Maximal Filtered Graph. See TMFG

algorithm A string indicating the algorithm to use or a function from igraph. Defaults to "walktrap". Current options are:
  - walktrap Computes the Walktrap algorithm using cluster_walktrap
  - leiden Computes the Leiden algorithm using cluster_louvain

algorithm.args List. A list of additional arguments for cluster_walktrap or cluster_leiden. Options are:
  - steps Number of steps used in the Walktrap algorithm. Defaults to c(3:8)
  - leiden Resolution parameter used in the Leiden algorithm. Defaults to seq(0, 2, .001). Higher values lead to smaller communities, lower values lead to larger communities

Value

Returns a list containing:

EGA The EGA output for the best fitting model
steps The number of steps used in the best fitting model from the cluster_walktrap algorithm
resolution_parameter The resolution parameter used in the best fitting model from the cluster_leiden algorithm
EntropyFit The tefi Index for the unique solutions given the range of steps (vector names represent the number of steps)
LowestEntropyFit The lowest value for the tefi Index
Author(s)
Hudson Golino <hfg9s at virginia.edu> and Alexander P. Christensen <alexpaulchristensen@gmail.com>

References

# Entropy fit measures

# Simulation for EGA.fit

# Leiden algorithm

# Walktrap algorithm

See Also

`bootEGA` to investigate the stability of EGA’s estimation via bootstrap, `EGA` to estimate the number of dimensions of an instrument using EGA, and `CFA` to verify the fit of the structure suggested by EGA using confirmatory factor analysis.

Examples

```r
# Load data
wmt <- wmt2[,7:24]

# Estimate EGA
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
ega.wmt <- EGA(data = wmt, plot.type = "qgraph")

# Estimate optimal EGA
fit.wmt <- EGA.fit(data = wmt)

# Plot optimal fit
plot(fit.wmt$EGA, plot.type = "qgraph")

# Compare with CFA
cfa.ega <- CFA(ega.wmt, estimator = "WLSMV", data = wmt)
cfa.fit <- CFA(fit.wmt$EGA, estimator = "WLSMV", data = wmt)
lavaan::lavTestLRT(cfa.ega$fit, cfa.fit$fit, method = "satorra.bentler.2001")
```
ega.wmt  

**EGA WMT-2 Data**

**Description**

EGA Network of wmt2 Data  
An EGA using the "glasso" model of the Wiener Matrizen-Test 2 (WMT-2)

**Usage**

data(ega.wmt)

data(ega.wmt)

**Format**

A 17 x 17 adjacency matrix  
A 17 x 17 adjacency matrix

**Details**

An EGA using the "glasso" model of the Wiener Matrizen-Test 2 (WMT-2)

**Examples**

data("ega.wmt")  
data("ega.wmt")

---

**Embed**  

**Time-delay Embedding**

**Description**

Reorganizes an individual’s observed time series into an embedded matrix. The embedded matrix is  
constructed with replicates of an individual time series that are offset from each other in time. The  
function requires two parameters, one that specifies the number of observations to be used (i.e. the  
number of embedded dimensions) and the other that specifies the number of observations to offset  
successive embeddings.

**Usage**

Embed(x, E, tau)
**entropyFit**

**Description**

 Computes the fit of a dimensionality structure using empirical entropy. Lower values suggest better fit of a structure to the data.

**Usage**

```r
test <- entropyFit(data, structure)
```

** Arguments **

- `x` Vector. An observed time series to be reorganized into a time-delayed embedded matrix.
- `E` Integer. Number of embedded dimensions or the number of observations to be used. For example, an "E = 5" will generate a matrix with five columns, meaning that five consecutive observations are used to create each row of the embedded matrix.
- `tau` Integer. Number of observations to offset successive embeddings. A tau of one uses adjacent observations. Default is "tau = 1".

**Value**

Returns a matrix containing the embedded matrix.

**Author(s)**

Pascal Deboeck <pascal.deboeck at psych.utah.edu>

**References**


**Examples**

```r
# A time series with 8 time points
tseries <- 49:56
embed.tseries <- Embed(tseries, E = 4, tau = 1)
```

**entropyFit**

**Entropy Fit Index**

Computes the fit of a dimensionality structure using empirical entropy. Lower values suggest better fit of a structure to the data.
entropyFit

Arguments

- **data**: Matrix or data frame. Contains variables to be used in the analysis
- **structure**: A vector representing the structure (numbers or labels for each item). Can be theoretical factors or the structure detected by EGA

Value

Returns a list containing:

- **Total.Correlation**: The total correlation of the dataset
- **Total.Correlation.MM**: Miller-Madow correction for the total correlation of the dataset
- **Entropy.Fit**: The Entropy Fit Index
- **Entropy.Fit.MM**: Miller-Madow correction for the Entropy Fit Index
- **Average.Entropy**: The average entropy of the dataset

Author(s)

Hudson F. Golino <hfg9s@virginia.edu>, Alexander P. Christensen <alexpaulchristensen@gmail.com> and Robert Moulder <rgm4fd@virginia.edu>

References


See Also

EGA to estimate the number of dimensions of an instrument using EGA and CFA to verify the fit of the structure suggested by EGA using confirmatory factor analysis.

Examples

```r
# Load data
wmt <- wmt2[,7:24]

# Estimate EGA model
ega.wmt <- EGA(data = wmt, model = "glasso", plot.EGA = FALSE)

# Compute entropy indices
entropyFit(data = wmt, structure = ega.wmt$wc)
```
ergoInfo

**Ergodicity Information Index**

**Description**
Computes the Ergodicity Information Index

**Usage**

```r
ergoInfo(data, use = c("edge.list", "weights"))
```

**Arguments**

- **data**
  A dynEGA.ind.pop object

- **use**
  Character. A string indicating what network element will be used to compute the algorithm complexity, the list of edges or the weights of the network. Defaults to use = "edge.list". Current options are:
  - `edge.list` Calculates the algorithm complexity using the list of edges.
  - `weights` Calculates the algorithm complexity using the weights of the network.

**Value**
Returns a list containing:

- **PrimeWeight**
  The prime-weight encoding of the individual networks

- **PrimeWeight.pop**
  The prime-weight encoding of the population network

- **Kcomp**
  The Kolmogorov complexity of the prime-weight encoded individual networks

- **Kcomp.pop**
  The Kolmogorov complexity of the prime-weight encoded population network

- **EII**
  The Ergodicity Information Index

**Author(s)**

Hudson Golino <hfg9s at virginia.edu>
Description

Estimates the derivatives of a time series using generalized local linear approximation (GLLA). GLLA is a filtering method for estimating derivatives from data that uses time delay embedding and a variant of Savitzky-Golay filtering to accomplish the task.

Usage

\texttt{glla(x, n.embed, tau, delta, order)}

Arguments

- **x**: Vector. An observed time series.
- **n.embed**: Integer. Number of embedded dimensions (the number of observations to be used in the \texttt{Embed} function).
- **tau**: Integer. Number of observations to offset successive embeddings in the \texttt{Embed} function. A tau of one uses adjacent observations. Default is "\texttt{tau = 1}".
- **delta**: Integer. The time between successive observations in the time series. Default is "\texttt{delta = 1}".
- **order**: Integer. The maximum order of the derivative to be estimated. For example, "\texttt{order = 2}" will return a matrix with three columns with the estimates of the observed scores and the first and second derivative for each row of the embedded matrix (i.e. the reorganization of the time series implemented via the \texttt{Embed} function).

Value

Returns a matrix containing \(n\) columns, in which \(n\) is one plus the maximum order of the derivatives to be estimated via generalized local linear approximation.

Author(s)

Hudson Golino <hfg9s at virginia.edu>

References


**Examples**

```r
# A time series with 8 time points
tseries <- 49:56
deriv.tseries <- glla(tseries, n.embed = 4, tau = 1, delta = 1, order = 2)
```

---

**hierEGA**

*Hierarchical EGA*

**Description**

Estimates EGA using the lower-order solution of `cluster_louvain` to identify the lower-order dimensions and then uses factor or network loadings to estimate factor or network scores, which are used to estimate the higher-order dimensions.

**Usage**

```r
hierEGA(
  data,
  scores = c("factor", "network"),
  consensus.iter = 1000,
  consensus.method = c("highest_modularity", "most_common", "iterative", "lowest_tefi"),
  uni.method = c("expand", "LE"),
  corr = c("cor_auto", "pearson", "spearman"),
  model = c("glasso", "TMFG"),
  model.args = list(),
  algorithm = c("walktrap", "louvain"),
  algorithm.args = list(),
  plot.EGA = TRUE,
  plot.type = c("GGally", "qgraph"),
  plot.args = list(),
  verbose = TRUE
)
```

**Arguments**

- **data**  
  Matrix or data frame. Variables (down columns) only. Does not accept correlation matrices.

- **scores**  
  Character. How should scores for the higher-order structure be estimated? Defaults to "network" for network scores computed using the `net.scores` function. Set to "factor" for factor scores computed using `fa`. Factors are assumed to be correlated using the "oblimin" rotation. **NOTE**: Factor scores
use the number of communities from EGA. Estimated factor may not align with these communities. The plots using factor scores with higher order factors that may not completely map onto the lower order communities. Look at the `$hierarchical$higher_order$lower_loadings` to determine the composition of the lower order factors.

By default, both factor and network scores are computed and stored in the output. The selected option only appears in the main output (`$hierarchical$`).

`consensus.iter` Numeric. Number of iterations to perform in consensus clustering (see Lancichinetti & Fortunato, 2012). Defaults to 1000.

`consensus.method` Character. What consensus clustering method should be used? Defaults to “highest_modularity”. Current options are:
- `highest_modularity` Uses the community solution that achieves the highest modularity across iterations
- `most_common` Uses the community solution that is found the most across iterations
- `iterative` Identifies the most common community solutions across iterations and determines how often nodes appear in the same community together. A threshold of 0.30 is used to set low proportions to zero. This process repeats iteratively until all nodes have a proportion of 1 in the community solution.
- `lowest_tefi` Uses the community solution that achieves the lowest `tefi` across iterations

By default, all consensus.method options are computed and stored in the output. The selected method will be used to plot and appear in the main output (`$hierarchical$`).

`uni.method` Character. What unidimensionality method should be used? Defaults to “LE”. Current options are:
- `expand` Expands the correlation matrix with four variables correlated .50. If number of dimension returns 2 or less in check, then the data are unidimensional; otherwise, regular EGA with no matrix expansion is used. This is the method used in the Golino et al. (2020) Psychological Methods simulation.
- `LE` Applies the leading eigenvalue algorithm (`cluster_leading_eigen`) on the empirical correlation matrix. If the number of dimensions is 1, then the leading eigenvalue solution is used; otherwise, regular EGA is used. This is the final method used in the Christensen, Garrido, and Golino (2021) simulation.

`corr` Type of correlation matrix to compute. The default uses `cor_auto`. Current options are:
- `cor_auto` Computes the correlation matrix using the `cor_auto` function from `qgraph`.
- `pearson` Computes Pearson’s correlation coefficient using the pairwise complete observations via the `cor` function.
- **spearman** Computes Spearman’s correlation coefficient using the pairwise complete observations via the `cor` function.

**model** Character. A string indicating the method to use. Defaults to "glasso". Current options are:

- **glasso** Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter
- **TMFG** Estimates a Triangulated Maximally Filtered Graph

**model.args** List. A list of additional arguments for `EBICglasso.qgraph` or `TMFG` algorithm. A string indicating the algorithm to use or a function from `igraph` Defaults to "louvain". Current options are:

- **walktrap** Computes the Walktrap algorithm using `cluster_walktrap`  
- **louvain** Computes the Louvain algorithm using `cluster_louvain`

**algorithm.args** List. A list of additional arguments for `cluster_walktrap, cluster_louvain`, or some other community detection algorithm function (see examples)

**plot.EGA** Boolean. If TRUE, returns a plot of the network and its estimated dimensions. Defaults to TRUE

**plot.type** Character. Plot system to use. Current options are `qgraph` and `GGally`. Defaults to "GGally"

**plot.args** List. A list of additional arguments for the network plot. For `plot.type` = "qgraph":

- **vsize** Size of the nodes. Defaults to 6.

For `plot.type` = "GGally" (see `ggnet2` for full list of arguments):

- **vsize** Size of the nodes. Defaults to 6.
- **label.size** Size of the labels. Defaults to 5.
- **alpha** The level of transparency of the nodes, which might be a single value or a vector of values. Defaults to 0.7.
- **edge.alpha** The level of transparency of the edges, which might be a single value or a vector of values. Defaults to 0.4.
- **legend.names** A vector with names for each dimension
- **color.palette** The color palette for the nodes. For custom colors, enter HEX codes for each dimension in a vector. See `color_palette_EGA` for more details and examples

**verbose** Boolean. Should network estimation parameters be printed? Defaults to TRUE. Set to FALSE for no print out

**Value**

Returns a list of lists containing:

**Main Results**
hierarchical

The main results list containing:

- **lower_order** Lower order EGA results for the selected methods
- **higher_order** Higher order EGA results for the selected methods

If `plot.EGA = TRUE`, then:

- **lower_plot** Plot of the lower order results
- **higher_plot** Plot of the higher order results
- **hier_plot** Plot of the lower and higher order results together, side-by-side

Secondary Results

**lower_ega** A list containing the lower order EGA results. The $\$we$ does not contain valid results. Do not use its output.

**lower_wc** A list containing consensus clustering results:

- **highest_modularity** Community memberships based on the highest modularity across the `cluster_louvain` applications
- **most_common** Community memberships based on the most commonly found memberships across the `cluster_louvain` applications
- **iterative** Community memberships based on consensus clustering described by Lancichinetti & Fortunato (2012)
- **lowest_tefi** Community memberships based on the lowest `tefi` across the `cluster_louvain` applications
- **summary_table** A data frame summarizing the unique community solutions across the iterations. Down the columns indicate: number of dimensions (N_Dimensions), proportion of times each community solution was identified (Proportion), modularity of each community solution (Modularity), total entropy fit index of each community solution (tefi), and the memberships for each item. Across the rows indicate each unique community solution

**factor_results** A list containing higher order results based on factor scores. A list for each consensus.method is provided with their EGA results

**network_results** A list containing higher order results based on network scores. A list for each consensus.method is provided with their EGA results

Author(s)

Marcos Jimenez <marcosjnezquez@gmail.com>, Francisco J. Abad <fjose.abad@uam.es>, Eduardo Garcia-Garzon <egarcia@ucjc.edu>, Luis Eduardo Garrido <luisgarrido@pucmm.edu.do>, Alexander P. Christensen <alexpaulchristensen@gmail.com>, and Hudson Golino <hfg9s@virginia.edu>

References

**Examples**

# Obtain example data
data <- optimism

# hierEGA example (no plots)
opt.hier <- hierEGA(
data = optimism,
algorithm = "louvain",
plot.EGA = FALSE # no plots for CRAN check
)

## Not run:
# hierEGA example (plots)
opt.hier <- hierEGA(
data = optimism,
algorithm = "louvain"
)

# Save plots
ggplot2::ggsave(
  filename = "hierEGA_plot.png", # name of plot
  plot = opt.res$hierarchical$hier_plot, # plot to save
  height = 8, # figure height
  width = 10, # figure width
  dpi = 600 # dots per inch
)

## End(Not run)

---

**intelligenceBattery  Intelligence Data**

**Description**

A response matrix (n = 1152) of the International Cognitive Ability Resource (ICAR) intelligence battery developed by Condon and Revelle (2016).

A response matrix (n = 1152) of the International Cognitive Ability Resource (ICAR) intelligence battery developed by Condon and Revelle (2016).

**Usage**

```r
data(intelligenceBattery)
data(intelligenceBattery)
```
Format
A 1185x125 response matrix
A 1185x125 response matrix

Examples
```r
data("intelligenceBattery")
data("intelligenceBattery")
```

---

**invariance**  
*Measurement Invariance of EGA Structure*

---

**Description**
Estimates metric invariance of EGA or specified structure

**Usage**
```r
invariance(
  data, groups,
  memberships = NULL,
  type = c("loadings"),
  iter = 500,
  ncores,
  ...
)
```

**Arguments**
- `data`  
  Matrix or data frame. Variables to be used in the analysis
- `groups`  
  Vector. Group membership corresponding to each case in data
- `memberships`  
  Vector. Node membership for each community or factor. Defaults to NULL. When NULL, EGA is used to compute node memberships
- `type`  
  Character. Type of measurement invariance to estimate. Only includes "loadings" at the moment
- `iter`  
  Numeric. Number of iterations to perform for the permutation. Defaults to 500
- `ncores`  
  Numeric. Number of cores to use in computing results. Defaults to `parallel::detectCores()` / 2 or half of your computer's processing power. Set to 1 to not use parallel computing
  If you're unsure how many cores your computer has, then use the following code: `parallel::detectCores()`
- `...`  
  Arguments passed to EGA
Value

Returns a list containing:

- **memberships**: Original memberships provided in `memberships` or from `EGA` if NULL
- **EGA**: Original `EGA` results for the sample
- **groups**: 
  - EGA EGA results for each group
  - loadings Network loadings for each group
  - loadingsDifference Difference between the dominant loadings of each group
- **permutation**: 
  - groups Permutated groups across iterations
  - loadings Loadings for each group for each permutation
  - loadingsDifference Difference between the dominant loadings of each group for each permutation
- **results**: Data frame of the results (which are printed)

Author(s)

Laura Jamison <lj5yn@virginia.edu>, Alexander P. Christensen <alexpaulchristensen@gmail.com>, and Hudson F. Golino <hfg9s@virginia.edu>

Examples

```r
# Load data
wmt <- wmt2[-1, 7:24]

# Groups
groups <- rep(1:2, each = nrow(wmt) / 2)

# Measurement invariance
results <- invariance(wmt, groups, ncores = 2)
```

Description

Based on the `bootEGA` results, this function computes and plots the number of times an item (variable) is estimated in the same factor/dimension as originally estimated by `EGA` (item.replication). The output also contains each item's replication frequency (i.e., proportion of bootstraps that an item appeared in each dimension; item.dim.rep) as well as the average network loading for each item in each dimension (item.loadings).
Usage

itemStability(bootega.obj, IS.plot = TRUE, structure = NULL, ...)

Arguments

bootega.obj  A bootEGA object
IS.plot Should the plot be produced for item.replication? If TRUE, then a plot for the item.replication output will be produced. Defaults to TRUE
structure User specified dimensionality structure.
... Additional arguments. Used for deprecated arguments from previous versions of itemStability

Value

Returns a list containing:

membership A list containing:
  • empirical The empirical memberships from the empirical EGA result
  • unique The unique dimensions from the empirical EGA result
  • bootstrap The memberships from the replicate samples in the bootEGA results

item.stability A list containing:
  • empirical.dimensions The proportion of times each item replicated within the empirical EGA defined dimension. This EGA result is defined using the input from bootEGA
  • all.dimensions The proportion of times each item replicated in each of the empirical EGA defined dimensions. This EGA result is defined using the input from bootEGA

plot A plot of the number of times each item replicated within the empirical EGA defined dimension.
mean.loadings Matrix of the average standardized network loading (computed using net.loads) for each item in each dimension

Author(s)

Hudson Golino <hfg9s at virginia.edu> and Alexander P. Christensen <alexpaulchristensen@gmail.com>

References


See Also

EGA to estimate the number of dimensions of an instrument using EGA and CFA to verify the fit of the structure suggested by EGA using confirmatory factor analysis.

Examples

```r
# Load data
wmt <- wmt[,7:24]

## Not run: # Estimate EGA network
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
ega.wmt <- EGA(data = wmt, model = "glasso", plot.type = "qgraph")

# Estimate dimension stability
boot.wmt <- bootEGA(data = wmt, iter = 100, typicalStructure = TRUE, plot.typicalStructure = TRUE, model = "glasso", plot.type = "qgraph", type = "parametric", ncores = 2)

## End(Not run)

# Estimate item stability statistics
res <- itemStability(boot.wmt)

# Changing plot features (ggplot2)
### qgraph Defaults
res$plot + 
  ggplot2::scale_color_manual(values = rainbow(max(res$membership$unique)))

### Pastel
res$plot + 
  ggplot2::scale_color_brewer(palette = "Pastel1")

# Changing Legend (ignore warnings)
res$plot + 
  ggplot2::scale_color_discrete(labels = "Intelligence")
```

---

**Loadings Comparison Test**

**Description**

An algorithm to identify whether data were generated from a factor or network model using factor and network loadings. The algorithm uses heuristics based on theory and simulation. These heuristics were then submitted to several deep learning neural networks with 240,000 samples per model with varying parameters.
Usage

```r
LCT(
  data,
  n,
  iter = 100,
  dynamic = FALSE,
  dynamic.args = list(n.embed = 4, tau = 1, delta = 1, use.derivatives = 1)
)
```

Arguments

- **data**: Matrix or data frame. A data frame with the variables to be used in the test or a correlation matrix. If the data used is a correlation matrix, the argument `n` will need to be specified.
- **n**: Integer. Sample size (if the data provided is a correlation matrix).
- **iter**: Integer. Number of replicate samples to be drawn from a multivariate normal distribution (uses `mvtnorm::mvrnorm`). Defaults to 100.
- **dynamic**: Boolean. Is the dataset a time series where rows are time points and columns are variables? Defaults to FALSE.
- **dynamic.args**: List. Arguments to be used in `dynEGA`. Defaults:
  - `n.embed`: Number of embeddings: 4
  - `tau`: Lag: 1
  - `delta`: Delta: 1
  - `use.derivatives`: Derivatives: 1

Value

Returns a list containing:

- **empirical**: Prediction of model based on empirical dataset only.
- **bootstrap**: Prediction of model based on means of the loadings across the bootstrap replicate samples.
- **proportion**: Proportions of models suggested across bootstraps.

Author(s)

Hudson F. Golino <hfg9s at virginia.edu> and Alexander P. Christensen <alexpaulchristensen at gmail.com>

References

Examples

# Compute LCT
LCT(data = wmt2[,7:24])

# Factor model
LCT(data = psychTools::bfi[,1:25])

# Dynamic LCT
LCT(sim.dynEGA[sim.dynEGA$ID == 1,1:20], dynamic = TRUE)

Description

Computes the Louvain community detection algorithm (Blondel et al., 2008)

Usage

louvain(A, method = c("modularity", "tefi"), resolution = 1, corr = NULL)

Arguments

A          Matrix or data frame. A network adjacency matrix
method     Character. Whether modularity or tefi should be used to optimize communities. Defaults to "modularity"
resolution Numeric. Resolution parameter for computing modularity. Defaults to 1. Values smaller than 1 favor larger communities; values larger than 1 favor smaller communities
corr       Matrix or data frame. Correlation matrix to be used when method = "tefi"

Details

This version was adapted from the Matlab code available here: https://perso.uclouvain.be/vincent.blondel/research/louvain.html. The code was adjusted to mirror the results of cluster_louvain. The Louvain algorithm's results can vary depending on node ordering. In this version, nodes are not shuffled so that consistent results can be achieved with the same node ordering. Results from cluster_louvain will shuffle nodes within the function and therefore will sometimes produce similar results and sometimes produce slightly different results. This version is based all in R and therefore is slower than the version in igraph.
Value

Returns a list containing:

- `wc`: A matrix of lower to higher order community membership detected in the network
- `modularity`: A vector of modularity values corresponding the rows of the `wc` matrix

Author(s)

Alexander P. Christensen <alexpaulchristensen@gmail.com> and Hudson Golino <hfg9s@virginia.edu>

References


Examples

```r
# Load data
dep <- depression[,24:44]

# Estimate correlations
corr <- qgraph::cor_auto(dep)

# Estimate network
net <- EBICglasso.qgraph(corr, n = nrow(dep))

# Estimate communities using modularity
louvain(net, method = "modularity")

# Estimate communities using tefi
louvain(net, method = "tefi", corr = corr)
```

---

**mctest.ergoInfo**

**Monte-Carlo Test for the Ergodicity Information Index**

Description

Computes a Monte-Carlo Test for the Ergodicity Information Index, comparing the empirical Ergodicity Information index to values obtained in a Monte-Carlo simulation in which all individuals have a similar latent structure. The p-values in the Monte-Carlo test can be calculated as \((\text{sum}(\text{EII} >= \text{MC.EII})+1)/(\text{iter}+1)\) and as \((\text{sum}(\text{EII} <= \text{MC.EII})+1)/(\text{iter}+1)\), where EII is the empirical Ergodicity Information Index, MC.EII is the values of the Ergodicity Information Index obtained in the simulation, and iter is the number of random samples generated in the simulation. The two-sided p-value is computed as two times the lowest p-value. In the Monte-Carlo Test for the Ergodicity Information Index, the null hypothesis is that the empirical value of EII is equal to the Monte-Carlo value of EII obtained in multiple individuals with a similar latent structure. Small
values of \( p \) indicate that is very unlikely to obtain an EII as large as the one obtained in the empirical sample if the null hypothesis is true, thus there is convincing evidence that the empirical Ergodicity Information Index is different than it could be expected if all individuals had a similar latent structure, conditioned on the parameters used to simulate the data.

**Usage**

```r
mctest.ergoInfo(
  iter,
  N,
  EII,
  use,
  variab,
  timep,
  nfact,
  error,
  dfm,
  loadings,
  autoreg,
  crossreg,
  var.shock,
  cov.shock,
  embed,
  tau,
  delta,
  derivatives,
  model,
  model.args = list(),
  algorithm = c("walktrap", "louvain"),
  algorithm.args = list(),
  corr,
  ncores,
  ...
)
```

**Arguments**

- **iter**
  - Numeric integer. Number of random samples to generate in the Monte-Carlo simulation. At least 500 is recommended.

- **N**
  - Numeric integer. Number of individuals to simulate data from, using the `simDFM` function.

- **EII**
  - Numeric. Empirical Ergodicity Information Index obtained via the `ergoInfo` function.

- **use**
  - Character. A string indicating what network element will be used to compute the algorithm complexity in the `ergoInfo` function, the list of edges or the weights of the network. Defaults to `use = "edge.list"`. Current options are:
    - `edge.list` Calculates the algorithm complexity using the list of edges.
• weights Calculates the algorithm complexity using the weights of the network.

variab Number of variables per factor.
timep Number of time points.
nfact Number of factors.
error Value to be used to construct a diagonal matrix Q. This matrix is p x p covariance matrix Q that will generate random errors following a multivariate normal distribution with mean zeros. The value provided is squared before constructing Q.
dfm A string indicating the dynamical factor model to use. Defaults to "DAFS". Current options are:
  • DAFS Simulates data using the direct autoregressive factor score model. This is the default method
  • RandomWalk Simulates data using a dynamic factor model with random walk factor scores.
loadings Magnitude of the loadings.
autoreg Magnitude of the autoregression coefficients. Default is "autoreg = 0.8".
crossreg Magnitude of the cross-regression coefficients. Default is "crossreg = 0.1".
var.shock Magnitude of the random shock variance. Default is "var.shock = 0.18".
cov.shock Magnitude of the random shock covariance. Default is "cov.shock = 0.36".
embed Integer. Number of embedded dimensions (the number of observations to be used in the Embed function). For example, an "embed = 5" will use five observations to estimate a single derivative. Defaults to embed = 5.
tau Integer. Number of observations to offset successive embeddings in the Embed function. A tau of one uses adjacent observations. Default is "tau = 1".
delta Integer. The time between successive observations in the time series. Default is "delta = 1".
derivatives Integer. The order of the derivative to be used in the EGA procedure. Default to 1.
model Character. A string indicating the method to use. Defaults to glasso. Current options are:
  • glasso Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter. This is the default method
  • TMFG Estimates a Triangulated Maximally Filtered Graph
model.args List. A list of additional arguments for EBIClasso.qgraph or TMFG algorithm A string indicating the algorithm to use or a function from igraph
Current options are:
  • walktrap Computes the Walktrap algorithm using cluster_walktrap
  • louvain Computes the Walktrap algorithm using cluster_louvain
mctest.ergoInfo

algorithm.args
List. A list of additional arguments for \texttt{cluster_walktrap, cluster_louvain}, or some other community detection algorithm function (see examples)

corr
Type of correlation matrix to compute. The default uses \texttt{cor_auto}. Current options are:
  - \texttt{cor_auto} Computes the correlation matrix using the \texttt{cor_auto} function from \texttt{qgraph}.
  - \texttt{pearson} Computes Pearson’s correlation coefficient using the pairwise complete observations via the \texttt{cor} function.
  - \texttt{spearman} Computes Spearman’s correlation coefficient using the pairwise complete observations via the \texttt{cor} function.

ncores
Numeric. Number of cores to use in computing results. Defaults to parallel::detectCores() / 2 or half of your computer’s processing power. Set to 1 to not use parallel computing. Recommended to use maximum number of cores minus one
If you’re unsure how many cores your computer has, then use the following code: parallel:::detectCores()

... Additional arguments. Used for deprecated arguments from previous versions of EGA

Value
Returns a list containing:

mc.ergoInfo The values of the Ergodicity Information Index obtained in the Monte-Carlo Simulation
p.value.twosided The p-value of the Monte-Carlo test for the Ergodicity Information Index. The null hypothesis is that the empirical Ergodicity Information index is equal to the expected value of the EII if the all individuals had similar latent structures.
effect Indicates whether the empirical EII is greater or less then the Monte-Carlo obtained EII.
plot.dist Histogram of the bootstrapped ergodicity information index

Author(s)
Hudson Golino <hfg9s@virginia.edu>

Examples

```r
# Not run:
\donttest{
  dyn1 <- dynEGA.ind.pop(data = sim.dynEGA, n.embed = 5, tau = 1,
                          delta = 1, id = 21, group = 22, use.derivatives = 1,
                          model = "glasso", ncores = 2, corr = "pearson")

  eii1 <- ergoInfo(data = dyn1)$EII

  dist.ergoinfo <- mctest.ergoInfo(iter = 10, N = 10, EII = eii1,
```
methods.section

Automated Methods Section for EGAnet Objects

Description

This function accepts EGAnet objects and generates a Methods section for your analysis. The output is an HTML page containing the descriptions of the methods and parameters as well as a Reference section for appropriate citation.

Usage

methods.section(...,
  stats = c("net.loads", "net.scores", "dimensionStability", "itemStability")
)

Arguments

... EGAnet objects. Available methods (more methods will be added soon!):
  • EGA Exploratory graph analysis
  • bootEGA Bootstrap exploratory graph analysis
  • UVA Unique variable analysis

stats Methods section for statistics in EGAnet. Multiple statistics can be input. Available statistics:
  • net.loads Network loadings. Requires EGA object to be input
  • net.scores Network scores. Requires EGA object to be input
  • dimensionStability Structural consistency. Requires bootEGA object to be input
  • itemStability Item stability. Requires bootEGA object to be input

Value

Automated HTML Methods section in your default browser
Examples

# Estimate EGA
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
ega.wmt <- EGA(data = wmt2[,7:24], plot.type = "qgraph")

# EGA Methods section
if(interactive()){
  methods.section(ega.wmt)
}

# Estimate standardized network loadings
wmt.loads <- net.loads(ega.wmt)$std

# EGA Methods section with network loadings
if(interactive()){
  methods.section(ega.wmt, stats = "net.loads")
}

## Not run: # bootEGA example
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
boot.wmt <- bootEGA(data = wmt2[,7:24], iter = 500, plot.type = "qgraph",
type = "parametric", ncores = 2)

## End(Not run)
# EGA and bootEGA Methods section
if(interactive()){
  methods.section(ega.wmt, boot.wmt)
}

# Estimate structural consistency
sc.wmt <- dimensionStability(boot.wmt)

# EGA and bootEGA Methods section with structural consistency and item stability
if(interactive()){
  methods.section(boot.wmt, stats = c("dimensionStability", "itemStability"))
}

# EGA with network loadings and
# bootEGA Methods section with structural consistency and item stability
if(interactive()){
  methods.section(ega.wmt, boot.wmt, stats = c("net.loads", "dimensionStability", "itemStability"))
}
Description

Computes the between- and within-community strength of each item for each community. This function uses the comcat and stable functions to calculate the between- and within-community strength of each item, respectively.

Usage

net.loads(A, wc, pos.manifold = FALSE, min.load = 0, plot.NL = FALSE)

Arguments

A Matrix, data frame, or EGA object. A network adjacency matrix
wc Numeric or character vector. A vector of community assignments. If input into A is an EGA object, then wc is automatically detected
pos.manifold Boolean. Should a positive manifold be applied (i.e., should all dimensions be positively correlated)? Defaults to FALSE. Set to TRUE for a positive manifold
min.load Numeric. Sets the minimum loading allowed in the standardized network loading matrix. Values equal or greater than the minimum loading are kept in the output. Values less than the minimum loading are removed. This matrix can be viewed using print() or summary(). Defaults to 0
plot.NL Boolean. Should proportional loadings be plotted? Defaults to FALSE. Set to TRUE for plot with pie charts visualizing the proportion of loading associated with each dimension

Details

Simulation studies have demonstrated that a node’s strength centrality is roughly equivalent to factor loadings (Christensen, Golino, & Silvia, 2019; Hallquist, Wright, & Molenaar, in press). Hallquist and colleagues (in press) found that node strength represented a combination of dominant and cross-factor loadings. This function computes each node’s strength within each specified dimension, providing a rough equivalent to factor loadings (including cross-loadings).

For more details, type vignette("Network_Scores")

Value

Returns a list containing:

unstd A matrix of the unstandardized within- and between-community strength values for each node
std A matrix of the standardized within- and between-community strength values for each node
minLoad The minimum loading to appear in summary of network loadings. Use print() or summary() to view
plot A qgraph plot of the network loadings. Use plot to view
net.scores

Author(s)
Alexander P. Christensen <alexpaulchristensen@gmail.com> and Hudson Golino <hfg9s at virginia.edu>

References

Examples

```r
# Load data
wmt <- wmt2[,7:24]

## Not run:
# Estimate EGA
ega.wmt <- EGA(wmt)

## End(Not run)

# Network loadings
net.loads(ega.wmt)
```

net.scores

<table>
<thead>
<tr>
<th>Network Scores</th>
</tr>
</thead>
</table>

Description
This function computes network scores computed based on each node’s strength within each community (i.e., factor) in the network (see net.loads). These values are used as network “factor loadings” for the weights of each item. Notably, network analysis allows nodes to contribution to more than one community. These loadings are considered in the network scores. In addition, if the construct is a hierarchy (e.g., personality questionnaire; items in facet scales in a trait domain), then an overall score can be computed (see argument global). An important difference is that the network scores account for cross-loadings in their estimation of scores.

Usage
net.scores(data, A, wc, global = FALSE, impute, ...)
Arguments

data  Matrix or data frame. Must be a dataset
A     Matrix, data frame, or EGA object. An adjacency matrix of network data
wc    Numeric. A vector of community assignments. Not necessary if an EGA object
       is input for argument A
global Boolean. Should general network loadings be computed in scores? Defaults to
       FALSE. If there is more than one dimension and there is theoretically one global
       dimension, then general loadings of the dimensions onto the global dimension
       can be included in the weighted scores
impute Character. In the presence of missing data, imputation can be implemented. Currently, three
       options are available:
       • none No imputation is performed. This is the default.
       • mean The "mean" value of the columns are used to replace the missing data.
       • median The "median" value of the columns are used to replace the missing
         data.
... Additional arguments for EGA

Details

For more details, type vignette("Network_Scores")

Value

Returns a list containing:

unstd.scores  The unstandardized network scores for each participant and community (including
               the overall score)
std.scores     The standardized network scores for each participant and community (including
               the overall score)
commCor        Partial correlations between the specified or identified communities
loads          Standardized network loadings for each item in each dimension (computed using
               net.loads)

Author(s)

Alexander P. Christensen <alexpaulchristensen@gmail.com> and Hudson F. Golino <hfg9s at vir-
ginia.edu>

References

Research Methods*, 53, 1563-1580.

Christensen, A. P., Golino, H., & Silvia, P. J. (2020). A psychometric network perspective on the
validity and validation of personality trait questionnaires. *European Journal of Personality*, 34,
1095-1108.

### Examples

```r
# Load data
wmt <- wmt2[,7:24]

## Not run:
# Estimate EGA
ega.wmt <- EGA(wmt)

## End(Not run)
# Network scores
net.scores(data = wmt, A = ega.wmt)
```

---

### network.descriptives

*Descriptive Statistics for Networks*

#### Description

Computes descriptive statistics for network models

#### Usage

```r
network.descriptives(network)
```

#### Arguments

- `network` Matrix, data frame, `qgraph`, or `EGA` object

#### Value

Numeric vector including:

- **Mean_weight**: The average of the edge weights in the network
- **SD_weight**: The standard deviation of the edge weights in the network
- **Min_weight**: The minimum of the edge weights in the network
- **Max_weight**: The minimum of the edge weights in the network
- **Density**: The density of the network
- **ASPL**: The average shortest path length (ASPL) of the network (computed as unweighted)
- **CC**: The clustering coefficient (CC) of the network (computed as unweighted)
network.descriptives

\textbf{swn.rand} \quad \text{Small-worldness measure based on random networks:}
\[
\text{swn.rand} = \frac{\text{ASPL}/\text{ASPL}_{\text{random}}}{\text{CC}/\text{CC}_{\text{random}}}
\]

\text{swn.rand} > 1 \text{ suggests the network is small-world}

\textbf{swn.HG} \quad \text{Small-worldness measure based on Humphries \& Gurney (2008):}
\[
\text{swn.HG} = \frac{\text{transitivity}/\text{transitivity}_{\text{random}}}{\text{ASPL}/\text{ASPL}_{\text{random}}}
\]

\text{swn.HG} > 1 \text{ suggests the network is small-world}

\textbf{swn.TJHBL} \quad \text{Small-worldness measure based on Telesford, Joyce, Hayasaka, Burdette, \& Laurienti (2011):}
\[
\text{swn.TJHBL} = \frac{\text{ASPL}_{\text{random}}/\text{ASPL}}{\text{CC}/\text{CC}_{\text{latice}}} - \frac{(\text{CC}/\text{CC}_{\text{latice}})}{\text{ASPL}_{\text{random}}/\text{ASPL}}
\]

\text{swn.TJHBL} \text{ near } 0 \text{ suggests the network is small-world, positive values suggest more random network characteristics, negative values suggest more lattice network characteristics}

\textbf{scale-free_R-sq} \quad \text{The R-squared fit of whether the degree distribution follows the power-law (many small degrees, few large degrees)}

\textbf{Author(s)}
Hudson Golino \texttt{<hfg9s at virginia.edu>} \text{and Alexander P. Christensen \texttt{<alexpaulchristensen@gmail.com>}}

\textbf{References}

\# swn.HG

\# swn.TJHBL

\# scale-free_R-sq

\textbf{Examples}

\# Load data
wmt <- wmt2[,7:24]

\# EGA example
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
ega.wmt <- EGA(data = wmt, plot.type = "qgraph")

\# Compute descriptives
network.descriptives(ega.wmt)
optimism

---

**Optimism Data**

Description

A response matrix (n = 282) containing responses to 10 items of the Revised Life Orientation Test (LOT-R), developed by Scheier, Carver, & Bridges (1994).

A response matrix (n = 282) containing responses to 10 items of the Revised Life Orientation Test (LOT-R), developed by Scheier, Carver, & Bridges (1994).

Usage

```r
data(optimism)
```

```r
data(optimism)
```

Format

A 282x10 response matrix

A 282x10 response matrix

References


Examples

```r
data("optimism")
```

```r
data("optimism")
```
plots

S3Methods for Plotting

Description

Plots for EGAnet objects

Usage

## S3 method for class 'bootEGA'
plot(x, plot.type = c("GGally", "qgraph"),
plot.args = list(), produce = TRUE, ...)

## S3 method for class 'CFA'
plot(x, layout = "spring", vsize = 6, ...)

## S3 method for class 'dynEGA'
plot(x, title = "", plot.type = c("GGally", "qgraph"),
plot.args = list(), produce = TRUE, ...)

## S3 method for class 'dynEGA.Groups'
plot(x, ncol, nrow, title = "", plot.type = c("GGally", "qgraph"),
plot.args = list(), produce = TRUE, ...)

## S3 method for class 'dynEGA.Individuals'
plot(x, title = "", id = NULL,
plot.type = c("GGally", "qgraph"), plot.args = list(), produce = TRUE, ...)

## S3 method for class 'EGA'
plot(x, title = "", plot.type = c("GGally", "qgraph"),
plot.args = list(), produce = TRUE, ...)

## S3 method for class 'NetLoads'
plot(x, ...)

Arguments

x Object from EGAnet package
plot.type Character. Plot system to use. Current options are qgraph and GGally. Defaults to "GGally".
plot.args List. A list of additional arguments for the network plot. For plot.type = "qgraph":
  • vsize Size of the nodes. Defaults to 6.
For plot.type = "GGally" (see ggnnet2 for full list of arguments):
  • vsize Size of the nodes. Defaults to 6.
• **label.size** Size of the labels. Defaults to 5.
• **alpha** The level of transparency of the nodes, which might be a single value or a vector of values. Defaults to 0.7.
• **edge.alpha** The level of transparency of the edges, which might be a single value or a vector of values. Defaults to 0.4.
• **legend.names** A vector with names for each dimension
• **color.palette** The color palette for the nodes. For custom colors, enter HEX codes for each dimension in a vector. See `color_palette_EGA` for more details and examples

**produce**

Boolean. This argument is used internally. Should plot be produced? Defaults to TRUE

... Arguments passed on to

• **qgraph** Functions: bootEGA, dynEGA, dynEGA.Groups, dynEGA.Individuals, EGA, and net.loads
• **semPaths** Functions: CFA

**vsize**

Numeric. Size of vertices in CFA plots. Defaults to 6

**layout**

Character. Layout of plot (see `semPaths`). Defaults to "spring"

**ncol**

Numeric. Number of columns

**nrow**

Numeric. Number of rows

**title**

Character. Title of the plot. Defaults to ""

**id**

Numeric. An integer or character indicating the ID of the individual to plot

---

**Prime Numbers through 100,000**

**Description**

Numeric vector of primes generated from the primes package. Used in the function [EGAnet]{ergoInfo}. Not for general use.

**Usage**

```r
data(prime.num)
```

Hudson Golino <hfg9s at virginia.edu> and Alexander P. Christensen <alexpaulchristensen@gmail.com>
Format

A 1185x24 response matrix
A 1185x24 response matrix

Examples

data("prime.num")
data("prime.num")

Description

Prints for EGAnet objects

Usage

## S3 method for class 'dynEGA'
print(x, ...)

## S3 method for class 'dynEGA.Groups'
print(x, ...)

## S3 method for class 'dynEGA.Individuals'
print(x, ...)

## S3 method for class 'EGA'
print(x, ...)

## S3 method for class 'NetLoads'
print(x, ...)

## S3 method for class 'invariance'
print(x, ...)

## S3 method for class 'hierEGA'
print(x, ...)

Arguments

x Object from EGAnet package
...
  Additional arguments
**residualEGA**

**Value**
Prints EGAnet object

**Author(s)**
Hudson Golino <hfg9s at virginia.edu> and Alexander P. Christensen <alexpaulchristensen@gmail.com>

---

**Description**

residualEGA Estimates the number of dimensions after controlling for wording effects. EGA is applied in the residual of a random intercept item factor model (RIIFA) with one method factor and one substantive factor.

**Usage**

```r
residualEGA(data, manifests, lat, negative.items, plot = TRUE)
```

**Arguments**

- `data` Matrix or data frame. Includes the variables to be used in the residualEGA analysis
- `manifests` Character vector. Vector indicating the names of the variables (items) to be used in the analysis.
- `lat` Numeric integer. Number of latent factors to be estimated. Only one substantive latent factor is recommended in the current version of the function.
- `negative.items` Numeric vector. A numeric vector indicating the column of the negative items.
- `plot` Boolean. If TRUE, returns a plot of the residualized network and its estimated dimensions. Defaults to TRUE

**Value**

Returns a list containing:

- `openMx.model` OpenMX model
- `openMx.result` OpenMX results
- `openMx.std.par` OpenMX standardized parameters
- `ResidualMatrix` Residual matrix
- `EGA.Residuals` Results of the residualized EGA
- `Fit` Fit metrics of the network structure, calculated using the ggmfit function of the qgraph package
- `WordLoads` Loadings of the wording effects
riEGA

Random-Intercept EGA

Description

Estimates the number of substantive dimensions after controlling for wording effects. EGA is applied to a residual correlation matrix after subtracting and random intercept factor with equal unstandardized loadings from all the regular and unrecoded reversed items in the database.

Usage

riEGA(
  data,
  n = NULL,
  uni.method = c("expand", "LE"),
  corr = c("cor_auto", "pearson", "spearman"),
  model = c("glasso", "TMFG"),
  model.args = list(),
  algorithm = c("walktrap", "louvain"),
  algorithm.args = list(),
  plot.EGA = TRUE,
  plot.type = c("GGally", "qgraph"),
  plot.args = list(),
  estimator = c("auto", "WLSMV", "MLR"),
riEGA

lavaan.args = list(),
   verbose = TRUE
)

Arguments

data Matrix or data frame. Variables (down columns) or correlation matrix. If the
input is a correlation matrix, then argument n (number of cases) is **required**. Variables **MUST** be unrecoded – reversed items should **remain** reversed

n Integer. Sample size if data provided is a correlation matrix

uni.method Character. What unidimensionality method should be used? Defaults to “LE”. Current options are:

  - expand Expands the correlation matrix with four variables correlated .50. If number of dimension returns 2 or less in check, then the data are unidimensional; otherwise, regular EGA with no matrix expansion is used. This is the method used in the Golino et al. (2020) *Psychological Methods* simulation.
  - LE Applies the leading eigenvalue algorithm (*cluster_leading_eigen*) on the empirical correlation matrix. If the number of dimensions is 1, then the leading eigenvalue solution is used; otherwise, regular EGA is used. This is the final method used in the Christensen, Garrido, and Golino (2021) simulation.

corr Type of correlation matrix to compute. The default uses *cor_auto*. Current options are:

  - cor_auto Computes the correlation matrix using the *cor_auto* function from *qgraph*.
  - pearsone Computes Pearson’s correlation coefficient using the pairwise complete observations via the *cor* function.
  - spearman Computes Spearman’s correlation coefficient using the pairwise complete observations via the *cor* function.

model Character. A string indicating the method to use. Defaults to "glasso". Current options are:

  - glasso Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter
  - TMFG Estimates a Triangulated Maximally Filtered Graph

model.args List. A list of additional arguments for *EBICglasso.qgraph* or *TMFG*

algorithm A string indicating the algorithm to use or a function from *igraph* Defaults to "walktrap". Current options are:

  - walktrap Computes the Walktrap algorithm using *cluster_walktrap*
  - louvain Computes the Louvain algorithm using *cluster_louvain*

algorithm.args List. A list of additional arguments for *cluster_walktrap*, *cluster_louvain*, or some other community detection algorithm function (see examples)
plot.EGA  Boolean. If TRUE, returns a plot of the network and its estimated dimensions. Defaults to TRUE.

plot.type  Character. Plot system to use. Current options are qgraph and GGally. Defaults to "GGally".

plot.args  List. A list of additional arguments for the network plot. For plot.type = "qgraph":

- vsize Size of the nodes. Defaults to 6.

For plot.type = "GGally" (see ggnet2 for full list of arguments):

- vsize Size of the nodes. Defaults to 6.
- label.size Size of the labels. Defaults to 5.
- alpha The level of transparency of the nodes, which might be a single value or a vector of values. Defaults to 0.7.
- edge.alpha The level of transparency of the edges, which might be a single value or a vector of values. Defaults to 0.4.
- legend.names A vector with names for each dimension
- color.palette The color palette for the nodes. For custom colors, enter HEX codes for each dimension in a vector. See color_palette_EGA for more details and examples

estimator  Character. Estimator to use for random-intercept model (see Estimators for more details). Defaults to "auto", which selects "MLR" for continuous data and "WLSMV" for mixed and categorical data. Data are considered continuous data if they have 6 or more categories (see Rhemtulla, Brosseau-Liard, & Savalei, 2012).

lavaan.args  List. If reduce.method = "latent", then lavaan’s cfa function will be used to create latent variables to reduce variables. Arguments should be input as a list. Some example arguments (see lavOptions for full details).

verbose  Boolean. Should network estimation parameters be printed? Defaults to TRUE. Set to FALSE for no print out.

Value
Returns a list containing:

EGA  Results from EGA

RI  A list containing information about the random-intercept model (if the model converged):

- fit The fit object for the random-intercept model using cfa
- lavaan.args The arguments used in cfa
- loadings Standardized loadings from the random-intercept model
- correlation Residual correlations after accounting for the random-intercept model

Author(s)
Alejandro Garcia-Pardina <alejandrogp97@gmail.com>, Francisco J. Abad <fjose.abad@uam.es>, Alexander P. Christensen <alexpaulchristensen@gmail.com>, Hudson Golino <hfg9s at virginia.edu>, Luis Eduardo Garrido <luisgarrido@pucmm.edu.do>, and Robert Moulder <rgm4fd@virginia.edu>
# Selection of CFA Estimator

**Examples**

```r
# Obtain example data
data <- optimism
## Not run:
# riEGA example
opt.res <- riEGA(data = optimism)
## End(Not run)
```

---

**sim.dynEGA Data**

**Description**

A simulated (multivariate time series) data with 20 variables, 200 individual observations, 50 time points per individual and 2 groups of individuals.

**Usage**

data(sim.dynEGA)

data(sim.dynEGA)

**Format**

A 10000x22 multivariate time series

**Examples**

data("sim.dynEGA")

data("sim.dynEGA")
Simulate data following a Dynamic Factor Model

Description
Function to simulate data following a dynamic factor model (DFM). Two DFMs are currently available: the direct autoregressive factor score model (Engle & Watson, 1981; Nesselroade, McArdle, Aggen, and Meyers, 2002) and the dynamic factor model with random walk factor scores.

Usage
simDFM(
  variab,  # Number of variables per factor.
  timep,  # Number of time points.
  nfact,  # Number of factors.
  error,  # Value to be used to construct a diagonal matrix Q. This matrix is p x p covariance matrix Q that will generate random errors following a multivariate normal distribution with mean zeros. The value provided is squared before constructing Q.
  dfm = c("DAFS", "RandomWalk"),  # A string indicating the dynamical factor model to use. Current options are:
  loadings,  # Magnitude of the loadings.
  autoreg,  # Magnitude of the autoregression coefficients.
  crossreg,  # Magnitude of the cross-regression coefficients.
  var.shock,  # Magnitude of the random shock variance.
  cov.shock,  # Magnitude of the random shock covariance
  burnin = 1000  # Number of n first samples to discard when computing the factor scores. Defaults to 1000.
)

Arguments

Variab | Number of variables per factor.
Timep | Number of time points.
Nfact | Number of factors.
Error | Value to be used to construct a diagonal matrix Q. This matrix is p x p covariance matrix Q that will generate random errors following a multivariate normal distribution with mean zeros. The value provided is squared before constructing Q.
Dfm | A string indicating the dynamical factor model to use. Current options are:
  - Dafs Simulates data using the direct autoregressive factor score model. This is the default method
  - RandomWalk Simulates data using a dynamic factor model with random walk factor scores.
Loadings | Magnitude of the loadings.
Autoreg | Magnitude of the autoregression coefficients.
Crossreg | Magnitude of the cross-regression coefficients.
Var.shock | Magnitude of the random shock variance.
Cov.shock | Magnitude of the random shock covariance
Burnin | Number of n first samples to discard when computing the factor scores. Defaults to 1000.
Author(s)

Hudson F. Golino <hfg9s at virginia.edu>

References


Examples

```r
## Not run:
donttest{
  # Estimate EGA network
data1 <- simDFM(variab = 5, timep = 50, nfact = 3, error = 0.05,
dfm = "DAFS", loadings = 0.7, autoreg = 0.8,
crossreg = 0.1, var.shock = 0.18,
cov.shock = 0.36, burnin = 1000)
}
## End(Not run)
```

### Summarys

**S3Methods for Summaries**

**Description**

Summaries for EGANet objects

**Usage**

```r
## S3 method for class 'CFA'
summary(object, ...)

## S3 method for class 'dynEGA'
summary(object, ...)

## S3 method for class 'dynEGA.Groups'
summary(object, ...)

## S3 method for class 'dynEGA.Individuals'
summary(object, ...)
```
### tefi

**Total Entropy Fit Index using Von Neumann's entropy (Quantum Information Theory) for correlation matrices**

**Description**

Computes the fit (TEFI) of a dimensionality structure using Von Neumann’s entropy when the input is a correlation matrix. Lower values suggest better fit of a structure to the data.

**Usage**

```r
tefi(data, structure)
```

**Arguments**

- `data`: A dataframe or correlation matrix
- `structure`: A vector representing the structure (numbers or labels for each item). Can be theoretical factors or the structure detected by EGA

**Value**

Returns a list containing:

- `VN.Entropy.Fit`: The Entropy Fit Index using Von Neumann’s entropy
- `Total.Correlation`: The total correlation of the dataset
- `Average.Entropy`: The average entropy of the dataset

**Arguments**

- `object`: Object from EGAnet package
- `...`: Additional arguments

**Value**

Summarizes EGAnet object

**Author(s)**

Hudson Golino <hfg9s at virginia.edu> and Alexander P. Christensen <alexpaulchristensen@gmail.com>
TMFG

Author(s)

Hudson Golino <chfg9s@virginia.edu>, Alexander P. Christensen <alexpaulchristensen@gmail.com>, and Robert Moulder <rgm4fd@virginia.edu>

References


See Also

EGA to estimate the number of dimensions of an instrument using EGA and CFA to verify the fit of the structure suggested by EGA using confirmatory factor analysis.

Examples

```r
# Load data
def <- def[, 7:24]

## Not run:
# Estimate EGA model
tmfg.def <- EGA(data = wmt, model = "glasso")

## End(Not run)

tmf <- tefi(data = tmfg.def$correlation, structure = tmfg.def$wc)
```

Description

Applies the Triangulated Maximally Filtered Graph (TMFG) filtering method (please see and cite Massara et al., 2016). The TMFG method uses a structural constraint that limits the number of zero-order correlations included in the network (3n - 6; where n is the number of variables). The TMFG algorithm begins by identifying four variables which have the largest sum of correlations to all other variables. Then, it iteratively adds each variable with the largest sum of three correlations to nodes already in the network until all variables have been added to the network. This structure can be associated with the inverse correlation matrix (i.e., precision matrix) to be turned into a GGM (i.e., partial correlation network) by using Local-Global Inversion Method (see Barfuss et al., 2016 for more details). See Details for more information on this network estimation method.
Usage

TMFG(cormat)

Arguments

cormat        A correlation matrix

Details

The TMFG method applies a structural constraint on the network, which restrains the network to retain a certain number of edges \((3n-6)\), where \(n\) is the number of nodes; Massara et al., 2016). The network is also composed of 3- and 4-node cliques (i.e., sets of connected nodes; a triangle and tetrahedron, respectively). The TMFG method constructs a network using zero-order correlations and the resulting network can be associated with the inverse covariance matrix (yielding a GGM; Barfuss, Massara, Di Matteo, & Aste, 2016). Notably, the TMFG can use any association measure and thus does not assume the data is multivariate normal.

Construction begins by forming a tetrahedron of the four nodes that have the highest sum of correlations that are greater than the average correlation in the correlation matrix. Next, the algorithm iteratively identifies the node that maximizes its sum of correlations to a connected set of three nodes (triangles) already included in the network and then adds that node to the network. The process is completed once every node is connected in the network. In this process, the network automatically generates what’s called a planar network. A planar network is a network that could be drawn on a sphere with no edges crossing (often, however, the networks are depicted with edges crossing; Tumminello, Aste, Di Matteo, & Mantegna, 2005).

Value

Returns a list containing:

\( A \)        The filtered adjacency matrix

separators     The separators (3-cliques) in the network

cliques        The cliques (4-cliques) in the network

Author(s)

Alexander Christensen <alexpaulchristensen@gmail.com>

References


Examples

# Pearson's correlation only for CRAN checks
A <- TMFG(cor(wmt2[,7:24]))$A

totalCor(data)

Arguments

data Matrix or data frame. Variables to be used in the analysis

Value

Returns a list containing:

Ind. Entropies Individual entropies for each variable
Joint. Entropy The joint entropy of the dataset
Total. Cor The total correlation of the dataset

Author(s)

Hudson F. Golino <hfg9s@virginia.edu>

References


Examples

# Compute total correlation
totalCor(wmt2[,7:24])
totalCorMat

Total Correlation Matrix

Description
Computes the pairwise total correlation for a dataset

Usage
```
totalCorMat(data)
```

Arguments
data

Matrix or data frame. Variables to be used in the analysis

Value
Returns a square matrix with pairwise total correlations

Author(s)
Hudson F. Golino <hfg9s at virginia.edu>

References


Examples
```
## Not run:
\donttest{
# Compute total correlation
totalCorMat(wmt2[,7:24])
}
## End(Not run)
**toy.example**  
*Toy Example Data*

**Description**

A simulated dataset with 2 factors, three items per factor and n = 500.

**Usage**

```r
data(toy.example)
```

**Format**

A 500x6 response matrix

**Examples**

```r
data("toy.example")
```

---

**UVA**  
*Unique Variable Analysis*

**Description**

Identifies redundant variables in a multivariate dataset using a number of different association methods and types of significance values (see Christensen, Garrido, & Golino, 2020 for more details)

**Usage**

```r
UVA(
  data,
  n = NULL,
  model = c("glasso", "TMFG"),
  corr = c("cor_auto", "pearson", "spearman"),
  method = c("cor", "pcor", "wTO"),
  type = c("adapt", "alpha", "threshold"),
  sig,
  key = NULL,
  reduce = TRUE,
  auto = TRUE,
  label_latent = TRUE,
  reduce.method = c("latent", "remove", "sum"),
  lavaan.args = list(),
  adhoc = TRUE,
  plot.redundancy = FALSE,
  plot.args = list()
)
```
Arguments

data  Matrix or data frame. Input can either be data or a correlation matrix
n    Numeric. If input in data is a correlation matrix, then sample size is required. Defaults to NULL
model Character. A string indicating the method to use. Current options are:
    • glasso Estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select optimal regularization parameter. This is the default method
    • TMFG Estimates a Triangulated Maximally Filtered Graph
corr Type of correlation matrix to compute. The default uses cor_auto. Current options are:
    • cor_auto Computes the correlation matrix using the cor_auto function from qgraph.
    • pearson Computes Pearson’s correlation coefficient using the pairwise complete observations via the cor function.
    • spearman Computes Spearman’s correlation coefficient using the pairwise complete observations via the cor function.
method Character. Computes weighted topological overlap ("wTO" using EBICglasso), partial correlations ("pcor"), or correlations ("cor") Defaults to "wTO"
type Character. Type of significance. Computes significance using the standard p-value ("alpha"), adaptive alpha p-value (adapt.a), or some threshold "threshold". Defaults to "threshold"
sig    Numeric. p-value for significance of overlap (defaults to .05). Defaults for "threshold" for each method:
        • "wTO" .25
        • "pcor" .35
        • "cor" .50
key    Character vector. A vector with variable descriptions that correspond to the order of variables input into data. Defaults to NULL or the column names of data
reduce Boolean. Should redundancy reduction be performed? Defaults to TRUE. Set to FALSE for redundancy analysis only
auto Boolean. Should redundancy reduction be automated? Defaults to TRUE. Set to FALSE for manual selection
label_latent Boolean. Should latent variables be labelled? Defaults to TRUE. Set to FALSE for arbitrary labelling (i.e., "LV_")
reduce.method Character. How should data be reduced? Defaults to "latent"
    • "latent" Redundant variables will be combined into a latent variable
    • "remove" All but one redundant variable will be removed
    • "sum" Redundant variables are combined by summing across cases (rows)
lavaan.args
List. If reduce.method = "latent", then lavaan's cfa function will be used to create latent variables to reduce variables. Arguments should be input as a list. Some example arguments (see lavOptions for full details):

- estimator Estimator to use for latent variables (see Estimators) for more details. Defaults to "MLR" for continuous data and "WLSMV" for mixed and categorical data. Data are considered continuous data if they have 6 or more categories (see Rhemtulla, Brosseau-Liard, & Savalei, 2012)
- missing How missing data should be handled. Defaults to "fiml"
- std.lv If TRUE, the metric of each latent variable is determined by fixing their (residual) variances to 1.0. If FALSE, the metric of each latent variable is determined by fixing the factor loading of the first indicator to 1.0. If there are multiple groups, std.lv = TRUE and "loadings" is included in the group label argument, then only the latent variances i of the first group will be fixed to 1.0, while the latent variances of other groups are set free. Defaults to TRUE

adhoc
Boolean. Should adhoc check of redundancies be performed? Defaults to TRUE. If TRUE, adhoc check will run the redundancy analysis on the reduced variable set to determine if there are any remaining redundancies. This check is performed with the arguments: method = "wTO", type = "threshold", and sig = .20. This check is based on Christensen, Garrido, and Golino’s (2020) simulation where these parameters were found to be the most conservative, demonstrating few false positives and false negatives

plot.redundancy
Boolean. Should redundancies be plotted in a network plot? Defaults to FALSE

plot.args
List. Arguments to be passed onto ggnet2. Defaults:

- vsize = 6 Changes node size
- alpha = 0.4 Changes transparency
- label.size = 5 Changes label size
- edge.alpha = 0.7 Changes edge transparency

Value
Returns a list:

redundancy A list containing several objects:

- redundant Vectors nested within the list corresponding to redundant nodes with the name of object in the list
- data Original data
- correlation Correlation matrix of original data
- weights Weights determine by weighted topological overlap, partial correlation, or zero-order correlation
- network If method = "wTO", then the network computed following EGA with EBICglasso network estimation
- plot If redundancy.plot = TRUE, then a plot of all redundancies found
- descriptives
- basic A vector containing the mean, standard deviation, median, median absolute deviation (MAD), 3 times the MAD, 6 times the MAD, minimum, maximum, and critical value for the overlap measure (i.e., weighted topological overlap, partial correlation, or threshold)
- centralTendency A matrix for all (absolute) non-zero values and their respective standard deviation from the mean and median absolute deviation from the median

* method Returns method argument
* type Returns type argument
* distribution If type != "threshold", then distribution that was used to determine significance

**reduced** If reduce = TRUE, then a list containing:
* data New data with redundant variables merged or removed
* merged A matrix containing the variables that were decided to be redundant with one another
* method Method used to perform redundancy reduction

**adhoc** If adhoc = TRUE, then the adhoc check containing the same objects as in the redundancy list object in the output

**Author(s)**

Alexander Christensen <alexpaulchristensen@gmail.com>

**References**

# Simulation using UVA

# Implementation of UVA (formally node.redundant)

# wTO measure

# Selection of CFA Estimator

**Examples**

# Select Five Factor Model personality items only
idx <- na.omit(match(gsub("-", ",", unlist(psychTools::spi.keys[1:5])), colnames(psychTools::spi)))
items <- psychTools::spi[,idx]
# Change names in redundancy output to each item's description
key.ind <- match(colnames(items), as.character(psychTools::spi.dictionary$item_id))
key <- as.character(psychTools::spi.dictionary$item[key.ind])

# Automated selection of redundant variables (default)
uva.results <- UVA(data = items, key = key)

# Manual selection of redundant variables
if(interactive()){
uva.results <- UVA(data = items, key = key, type = "adapt")
}

---

**vn.entropy**

**Entropy Fit Index using Von Neumann's entropy (Quantum Information Theory) for correlation matrices**

**Description**

Computes the fit of a dimensionality structure using Von Neumann's entropy when the input is a correlation matrix. Lower values suggest better fit of a structure to the data.

**Usage**

`vn.entropy(data, structure)`

**Arguments**

- `data` A datafram or a correlation matrix
- `structure` A vector representing the structure (numbers or labels for each item). Can be theoretical factors or the structure detected by EGA

**Value**

Returns a list containing:

- **VN.Entropy.Fit** The Entropy Fit Index using Von Neumann's entropy
- **Total.Correlation** The total correlation of the dataset
- **Average.Entropy** The average entropy of the dataset

**Author(s)**

Hudson Golino <hfg9s at virginia.edu>, Alexander P. Christensen <alexpaulchristensen@gmail.com>, and Robert Moulder <rgm4fd@virginia.edu>
References


See Also

EGA to estimate the number of dimensions of an instrument using EGA and CFA to verify the fit of the structure suggested by EGA using confirmatory factor analysis.

Examples

```r
# Load data
dep <- depression[,24:44]

# Estimate EGA
## plot.type = "qgraph" used for CRAN checks
## plot.type = "GGally" is the default
ga <- EGA(data = dep, model = "glasso", plot.type = "qgraph")

# Compute entropy indices
gn.entropy(data = ega$correlation, structure = ega$wc)
```

---

**wmt2**

**WMT-2 Data**

**Description**

A response matrix (n = 1185) of the Wiener Matrizen-Test 2 (WMT-2).

**Usage**

```r
data(wmt2)
```

**Format**

A 1185x24 response matrix

A 1185x24 response matrix
Examples

data("wmt2")

data("wmt2")
Index

* datasets
  boot.wmt, 8
  depression, 16
  dnn.weights, 19
  ega.wmt, 36
  intelligenceBattery, 45
  optimism, 63
  prime.num, 65
  sim.dynEGA, 71
  toy.example, 79
  wmt2, 84

boot.ergoInfo, 5
boot.wmt, 8
bootEGA, 3, 4, 8, 13, 17, 29, 32, 35, 47, 48, 56

CFA, 12, 12, 18, 29, 32, 35, 38, 49, 65, 68, 75, 84

cfa, 13, 70, 81
cluster_leading_eigen, 9, 27, 34, 42, 69
cluster_leiden, 34
cluster_louvain, 6, 10, 21, 23, 24, 28, 31, 34, 41, 43, 44, 51, 54, 55, 69
cluster_walktrap, 6, 10, 21, 23, 24, 28, 31, 33, 34, 43, 54, 55, 69

color_palette_EGA, 11, 14, 16, 28, 43, 65, 70
compare.EGA.plots, 15
cor, 6, 10, 21, 24, 27, 31, 34, 42, 43, 55, 69, 80
cor_auto, 6, 10, 21, 24, 27, 31, 34, 42, 55, 69, 80

depression, 16
dimensionStability, 4, 11, 17, 56
dnn.weights, 19
dynEGA, 3–5, 19, 50
dynEGA.ind.pop, 22

EBICglasso, 3, 24, 80, 81
EBICglasso.qgraph, 6, 10, 21, 23, 24, 26, 28–32, 34, 43, 54, 69

EGA, 3, 4, 7, 8, 11–14, 17, 18, 21, 24, 26, 28, 31, 33–36, 38, 41, 42, 44, 46–49, 55, 56, 58, 60, 61, 67, 68, 70, 74, 75, 81, 83, 84
EGA.estimate, 30
EGA.fit, 4, 5, 33
ega.wmt, 36
EGAnet, 15, 16, 56
EGAnet (EGAnet-package), 3
EGAnet-package, 3
Embed, 6, 20, 23, 36, 40, 54
entropyFit, 3, 4, 37
ergoInfo, 6, 39, 53

fa, 41

fitMeasures, 13

GGally, 10, 16, 28, 43, 64, 70
ggnet2, 10, 14, 16, 28, 43, 64, 70, 81
glasso, 24, 25
gll, 19, 22, 40

hierEGA, 41

igraph, 6, 10, 21, 23, 28, 31, 34, 43, 51, 54, 69
intelligenceBattery, 45
invariance, 46
itemStability, 4, 17, 47, 48, 56
lavaan, 70, 81
lavOptions, 12, 70, 81
LCT, 4, 49
louvain, 51

mctest.ergoInfo, 52
methods.section, 56
mvrnorm, 9

net.loads, 3, 4, 48, 56, 57, 59, 60
net.scores, 3, 41, 56, 59
network.descriptives, 61
INDEX

optimism, 63

plot.bootEGA (plots), 64
plot.CFA (plots), 64
plot.dynEGA (plots), 64
plot.EGA (plots), 64
plot.NetLoads (plots), 64
plots, 64
prime.num, 65
print.dynEGA (prints), 66
print.EGA (prints), 66
print.hierEGA (prints), 66
print.invariance (prints), 66
print.NetLoads (prints), 66
prints, 66

qgraph, 6, 10, 14, 16, 21, 24, 27, 28, 31, 34, 42, 43, 55, 58, 61, 64, 65, 67, 69, 70, 80

RCColorBrewer, 14
residualEGA, 67
riEGA, 68

semPaths, 13, 65
sim.dynEGA, 71
simDFM, 4, 53, 72
summary.CFA (summarys), 73
summary.dynEGA (summarys), 73
summary.EGA (summarys), 73
summary.NetLoads (summarys), 73
summarys, 73

tefi, 3–5, 33, 34, 42, 44, 51, 74
TMFG, 3, 6, 10, 21, 23, 26, 28–32, 34, 43, 54, 69, 75
totalCor, 77
totalCorMat, 78
toy.example, 79

UVA, 3, 4, 56, 79

vn.entropy, 4, 83

wi2net, 25
wmt2, 8, 36, 84