Package ‘alsi’

February 17, 2026

Type Package

Title Aggregated Latent Space Index for Multiple Correspondence Analysis

Version 0.1.2

Description Tools for stability-validated aggregation in multiple correspondence analysis (MCA). Implements parallel analysis for dimensionality assessment, bootstrap-based subspace stability diagnostics using Procrustes rotation and Tucker's congruence coefficients, and computation of the Aggregated Latent Space Index (ALSI). ALSI is a person-level summary measure derived from validated MCA dimensions that quantifies departure from independence along stable association directions in multivariate categorical data.

License GPL-3

Encoding UTF-8

LazyData true

Depends R (>= 4.0.0)

Imports stats, graphics, utils

Suggests readxl, openxlsx

RoxygenNote 7.2.0

NeedsCompilation no

Author Se-Kang Kim [aut, cre]

Maintainer Se-Kang Kim `<se-kang.kim@bcm.edu>`

Repository CRAN

Date/Publication 2026-02-17 16:00:02 UTC

Contents

`alsi` *Compute Aggregated Latent Space Index (ALSI)*

Description

Calculates ALSI as a variance-weighted Euclidean norm of row principal coordinates within a retained K-dimensional MCA subspace.

Usage

```
alsi(Fmat, eig, K)
```

Arguments

`Fmat`	Matrix of row principal coordinates ($N \times K$ or larger)
`eig`	Vector of eigenvalues (inertias)
`K`	Integer, number of dimensions to aggregate

Value

S3 object of class `alsi` containing:

`alpha`	Numeric vector of ALSI values (length N), representing each individual's variance-weighted distance from the centroid in the retained MCA subspace
`w`	Variance weights (length K), computed as the proportion of retained inertia for each dimension
`alpha_vec`	Aggregated direction vector (length K), equal to sqrt(w), used for projecting category coordinates
`K`	Number of dimensions used in aggregation

Examples

```
# Create example data
set.seed(123)
Fmat <- matrix(rnorm(100 * 4), nrow = 100, ncol = 4)
eig <- c(0.5, 0.3, 0.15, 0.05)

# Compute ALSI
a <- alsi(Fmat, eig, K = 3)
print(a)
hist(a$alpha, main = "Distribution of ALSI")
```

`alsi_workflow` *Example workflow using the ALSI package*

Description

Example workflow using the ALSI package

Usage

```
alsi_workflow(
  path,
  vars,
  B_pa = 2000,
  B_boot = 2000,
  q = 0.95,
  seed = 20260123
)
```

Arguments

`path`	Path to data file (.xlsx) or data frame
`vars`	Character vector of binary variable names
`B_pa`	Number of permutations for parallel analysis
`B_boot`	Number of bootstrap resamples
`q`	Quantile for parallel analysis
`seed`	Random seed

Value

List containing all analysis objects

Examples

```
# Complete workflow
results <- alsi_workflow(
  path = "ANR2.xlsx",
  vars = c("MDD", "DYS", "DEP", "PTSD", "OCD", "GAD", "ANX", "SOPH", "ADHD"),
  B_pa = 2000,
  B_boot = 2000
)

# Access components
results$pa       # Parallel analysis
results$boot     # Bootstrap stability
results$alsi     # ALSI values
```

`ANR2` *Example Eating Disorder Diagnostic Data*

Description

Individual-level binary diagnostic data for eating disorder patients, including nine psychiatric diagnoses and pre/post treatment measures.

Usage

```
ANR2
```

Format

A data frame with 1261 rows and 13 variables:

MDD	Major Depressive Disorder (0 = absent, 1 = present)
DYS	Dysthymia (0 = absent, 1 = present)
DEP	Depression (0 = absent, 1 = present)
PTSD	Post-Traumatic Stress Disorder (0 = absent, 1 = present)
OCD	Obsessive-Compulsive Disorder (0 = absent, 1 = present)
GAD	Generalized Anxiety Disorder (0 = absent, 1 = present)
ANX	Anxiety (0 = absent, 1 = present)
SOPH	Social Phobia (0 = absent, 1 = present)
ADHD	Attention-Deficit/Hyperactivity Disorder (0 = absent, 1 = present)
pre_bmi	Pre-treatment Body Mass Index (numeric)
post_bmi	Post-treatment Body Mass Index (numeric)
pre_EDI	Pre-treatment Eating Disorder Inventory score (numeric)
post_EDI	Post-treatment Eating Disorder Inventory score (numeric)

Source

Baylor College of Medicine eating disorder treatment program data. Data have been de-identified and anonymized for research purposes.

Examples

```
data(ANR2)

# Examine structure
str(ANR2)

# View diagnostic variables
vars <- c("MDD", "DYS", "DEP", "PTSD", "OCD", "GAD", "ANX", "SOPH", "ADHD")
head(ANR2[, vars])
```

```
# Check prevalence of diagnoses
colMeans(ANR2[, vars], na.rm = TRUE)
```

`mca_align` *Align MCA solution via Procrustes rotation with sign anchoring*

Description

Align MCA solution via Procrustes rotation with sign anchoring

Usage

```
mca_align(G, Gref)
```

Arguments

`G`	Matrix of category coordinates to align
`Gref`	Reference matrix of category coordinates

Value

List with aligned coordinates and rotation matrix

`mca_bootstrap` *Bootstrap-Based Subspace Stability Assessment*

Description

Evaluates reproducibility of retained MCA dimensions via bootstrap resampling. Quantifies stability using Procrustes principal angles (subspace-level) and Tucker's congruence coefficients (dimension-level).

Usage

```
mca_bootstrap(data, vars, K, B = 2000, seed = 20260123, verbose = TRUE)
```

Arguments

`data`	Data frame or path to .xlsx file
`vars`	Character vector of binary variable names
`K`	Integer, number of dimensions to retain and assess
`B`	Integer, number of bootstrap resamples (default: 2000)
`seed`	Integer, random seed for reproducibility
`verbose`	Logical, print progress messages

Value

S3 object of class `mca_bootstrap` containing:

`ref`	Reference MCA fit
`K`	Number of dimensions assessed
`B`	Number of bootstrap resamples
`angles`	Matrix of principal angles ($B \times K$)
`tucker`	Matrix of Tucker congruence coefficients ($B \times K$)
`angles_summary`	Summary statistics for angles
`tucker_summary`	Summary statistics for congruence

Examples

```
data(ANR2)
boot <- mca_bootstrap(ANR2, vars = names(ANR2), K = 3, B = 100)
print(boot)
plot(boot)
```

`mca_pa` *Parallel Analysis for MCA Dimensionality Assessment*

Description

Compares observed MCA eigenvalues against reference distributions from permuted data to identify statistically meaningful dimensions.

Usage

```
mca_pa(
  data,
  vars,
  B = 2000,
  q = 0.95,
  seed = 20260123,
  max_dims = 20,
  verbose = TRUE
)
```

Arguments

`data`	Data frame or path to .xlsx file
`vars`	Character vector of binary variable names
`B`	Integer, number of permutations (default: 2000)
`q`	Numeric, reference quantile for retention (default: 0.95)

`seed`	Integer, random seed for reproducibility
`max_dims`	Integer, maximum dimensions to display in plot
`verbose`	Logical, print progress messages

Value

S3 object of class `mca_pa` containing:

`eig_obs`	Observed eigenvalues from the MCA of the original data
`eig_q`	Reference quantiles from permutation distribution
`eig_perm`	Matrix of permutation eigenvalues (B x dimensions)
`K_star`	Suggested number of dimensions to retain (where observed > reference)
`fit`	MCA fit object (class `mca_fit`) from original data
`q`	Quantile threshold used for comparison
`B`	Number of permutations performed

Examples

```
# Using included ANR2 dataset
data(ANR2)
pa <- mca_pa(ANR2, vars = names(ANR2), B = 100)
print(pa$K_star)
```

`plot_category_projections`

Plot Category Projections in MCA Space

Description

Visualizes category coordinates in a 2D MCA subspace and optionally displays projections onto the aggregated ALSI direction.

Usage

```
plot_category_projections(
  fit,
  K,
  alpha_vec = NULL,
  dim_pair = c(1, 2),
  cex = 0.8,
  top_n = 15
)
```

Arguments

`fit`	MCA fit object (class `mca_fit`)
`K`	Number of dimensions in retained subspace
`alpha_vec`	Optional aggregated direction vector (from `alsi()`)
`dim_pair`	Integer vector of length 2, dimensions to plot (default: c(1,2))
`cex`	Character expansion for labels
`top_n`	Number of top categories to display by projection (default: 15)

Value

No return value, called for side effects. The function creates a scatter plot of category coordinates in the specified 2D subspace, with category labels displayed. If `alpha_vec` is provided, it also prints the top categories ranked by their absolute projection onto the ALSI direction to the console.

`plot_subspace_stability`

Plot Subspace Stability Diagnostics

Description

Creates diagnostic plots showing distributions of principal angles and Tucker congruence coefficients across bootstrap resamples.

Usage

```
plot_subspace_stability(boot_obj)
```

Arguments

`boot_obj`	Object of class `mca_bootstrap`

Value

No return value, called for side effects. The function creates a two-panel figure with: (1) boxplots of principal angles (left panel), showing the distribution of subspace similarity across bootstrap resamples for each dimension; and (2) boxplots of Tucker congruence coefficients (right panel), showing dimension-level replicability with reference lines at phi = 0.85 (good) and phi = 0.95 (excellent).

Index