

Package ‘MaxIntTools’

February 9, 2026

Type Package

Title Testing Maximal Interaction in Two-Mode Clustering via a
Permutation Based Procedure

Version 0.1.0

Description

Performs maximal interaction two-mode clustering, permutation tests, scree plots, and interaction visualizations for bicluster analysis. See

Ahmed et al. (2025) <[doi:10.17605/OSF.IO/AWGXB](https://doi.org/10.17605/OSF.IO/AWGXB)>, Ahmed et al. (2023) <[doi:10.1007/s00357-023-09434-2](https://doi.org/10.1007/s00357-023-09434-2)>,
Ahmed et al. (2021) <[doi:10.1007/s11634-021-00441-y](https://doi.org/10.1007/s11634-021-00441-y)>.

License GPL-3

Encoding UTF-8

Imports ggplot2, MASS, reshape, pracma

RoxygenNote 7.3.2

NeedsCompilation no

Author Zaheer Ahmed [aut, cre, cph],
Jan Schepers [ctb],
Uwe Ligges [ctb],
Alberto Cassese [ctb],
Gerard van Breukelen [ctb],
Katja Ickstadt [ctb],
Edmund Wascher [ctb]

Maintainer Zaheer Ahmed <ahmed@ifado.de>

Repository CRAN

Date/Publication 2026-02-09 13:10:21 UTC

Contents

MaxIntTools-package	2
extract_Gamma_EReMI	3
E_ReMI	4
Log_Likelihood_function_E_ReMI	6

Log_Likelihood_function_REMAXINT	8
Log_LR_Test_Statistic	9
MaxInt.heatmap.plot	11
MaxInteraction.plot	13
MaxInt_data_analysis	14
Permutation_Function	16
print.MaxInt_data_analysis	18
Randompartition_function	20
REMAXINT	21
screepplot.MaxInt_data_analysis	23
Unequal_Randompartition_Function	24
Update_column_clusters	26
Update_G_Omega	27
Update_row_clusters_E_ReMI	29
Update_row_clusters_REMAXINT	31

Index 33

MaxIntTools-package *MaxIntTools: Tools for Detecting and Understanding Maximal Interaction Two-Mode Clustering*

Description

MaxIntTools provides functions for performing maximal interaction two-mode clustering and related visualization techniques. It facilitates detection and understanding of latent interaction in complex two-mode data sets with applications in psychology, marketing, neuroscience, and other fields.

Details

The package includes the following key functions:

- `MaxInt_data_analysis`: Analysis two-mode data using this function.
- `MaxInt.Screepplot`: Plot all (P, Q) clustering combinations so that we can decide the optimal ones.
- `MaxInteraction.plot`: Plot the resulting of maximal interaction two-mode clustering for optimal choice.
- `MaxInt.heatmap.plot`: Generate heatmaps annotated with optimal clustering results.

These tools support exploratory data analysis and interpretation of complex interaction structures in two-mode data.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de) Maintainer: Zaheer Ahmed (ahmed@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[MaxInt_data_analysis](#), [screplot.MaxInt_data_analysis](#), [MaxInteraction.plot](#), [MaxInt.heatmap.plot](#)

extract_Gamma_EReMI *Extract Estimated Gamma Matrix for a Specific Cluster Combination*

Description

Extracts the estimated Gamma matrix corresponding to a specific (P, Q) cluster solution from the output of the E-ReMI algorithm.

Usage

```
extract_Gamma_EReMI(x, selection)
```

Arguments

x	An object returned by the EReMI() function, containing the list of estimated Gamma matrices for different cluster combinations.
selection	A character string or numeric vector specifying the cluster combination. If a character string, it should be in the form "(P, Q)", e.g., "(3, 4)". If numeric, it will be internally converted to the appropriate label.

Details

This function simplifies access to the estimated Gamma matrix for a user-specified combination of row (P) and column (Q) clusters, as identified in the output of the E-ReMI algorithm. It automatically reformats the input if the selection is provided in character form (e.g., "(3, 4)").

Value

A matrix of estimated Gamma values corresponding to the specified cluster combination.
If the cluster combination is invalid or not found, NULL is returned.

Note

Used internally by visualization and post-processing functions for E-ReMI output.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation-based alternative. *Available on OSF* (submitted to ADAC). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331. Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013.

See Also

[E_ReMI](#), [Permutation_Function](#), [MaxInt_data_analysis](#), [screepLOT.MaxInt_data_analysis](#), [MaxInteraction.plot](#)

E_ReMI

Two-Mode Clustering via the E_ReMI Method

Description

Performs row-by-column clustering using the E-ReMI (Extended REMAXINT) algorithm. It estimates latent cluster structures in rows and columns of a data matrix with unequal row cluster proportions using a REMAXINT-like procedure.

Usage

```
E_ReMI(DC, P, Q, Nruns, initR, initC, initG)
```

Arguments

DC	A doubly centered numeric matrix of size $I \times J$ representing the input data to be clustered.
P	An integer indicating the number of clusters for rows.
Q	An integer indicating the number of clusters for columns.
Nruns	The number of random initializations (runs) to perform to avoid local maxima. The solution with the highest likelihood is returned.

initR	An initial binary matrix of size $I \times P$ indicating the initial row cluster assignments.
initC	An initial binary matrix of size $J \times Q$ indicating the initial column cluster assignments.
initG	A $P \times Q$ matrix representing the initial estimate of the latent interaction between row and column clusters.

Details

The E_ReMI function implements an EM-style algorithm for biclustering, based on maximizing the likelihood of a Gaussian interaction model with unequal row cluster sizes. Starting from user-provided initial values for cluster assignments and the interaction matrix, the algorithm iteratively updates row clusters, column clusters, the bi-cluster interaction matrix, and estimates the proportions of row clusters (Ω) and residual variance. Unlike REMAXINT clustering models, E_ReMI accounts for unequal cluster sizes in expectation, allowing for more flexible modeling of heterogeneous block-structured data.

Value

A list containing:

BestR	Binary matrix ($I \times P$) of row cluster assignments.
BestC	Binary matrix ($J \times Q$) of column cluster assignments.
BestG	Estimated $P \times Q$ latent interaction matrix.
LL	Maximal log-likelihood of the fitted model.
Best_0	Estimated row cluster proportions (Ω), a vector of length P .

Note

This function is useful for clustering data with unequal row cluster sizes and estimating interaction structures between biclusters.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation-based alternative. *Available on OSF* (submitted to ADAC). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331. Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013.

See Also

[Permutation_Function](#), [Randompartition_function](#), [Unequal_Randompartition_Function](#), [Update_row_clusters_E_ReMI](#), [Update_column_clusters](#), [Update_G_Omega](#), [Log_Likelihood_function_E_ReMI](#), [pinv](#)

Examples

```
I <- 10
P <- 4
J <- 5
Q <- 2
Nruns <- 5
DC <- matrix(rnorm(I * J), I, J)
initR <- diag(1, nrow = I, ncol = P)
initC <- diag(1, nrow = J, ncol = Q)
R_inv <- pracma::pinv(t(initR) %*% initR)
C_inv <- pracma::pinv(t(initC) %*% initC)
initG <- R_inv %*% t(initR) %*% DC %*% initC %*% C_inv
initG <- as.matrix(initG)
result <- E_ReMI(DC, P, Q, Nruns, initR, initC, initG)
result
```

Log_Likelihood_function_E_ReMI

Log-Likelihood Calculation for E-ReMI Clustering

Description

Computes the log-likelihood of the observed data in the E-ReMI clustering framework.

Usage

```
Log_Likelihood_function_E_ReMI(DC, I, J, initM, initOmega_hat)
```

Arguments

DC	A numeric data matrix of dimension $I \times J$ representing the input data with double centering.
I	An integer indicating the number of rows (units) in the data matrix DC.
J	An integer indicating the number of columns (conditions or variables) in the data matrix DC.
initM	A numeric matrix of the same dimension as DC representing the reconstructed matrix based on current estimates of cluster assignments and latent interaction parameters. It is computed via random row partition matrix (R), random column partition matrix (C) and bi-cluster interaction matrix (G) and updated after each iteration.
initOmega_hat	A numeric vector of length equal to the number of row clusters, representing the estimated mixing proportions (probabilities) of row clusters.

Details

This function is used internally in the E-ReMI algorithm to monitor model fit by calculating the log-likelihood of the observed data under the current clustering configuration and model estimates. The error matrix is computed by subtracting the reconstructed matrix `ini tM` from the observed data DC, and the log-likelihood is derived from the Gaussian likelihood function with known variance.

Value

Returns a numeric scalar:

LL The log-likelihood value of the current E-ReMI model configuration.

Note

This function is intended for internal use in model monitoring and convergence checking during E-ReMI algorithm execution.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[E_ReMI](#), [Update_row_clusters_E_ReMI](#), [pinv](#)

Examples

```
I <- 10; P <- 4; J <- 5; Q <- 2
DC <- matrix(rnorm(I*J), I, J)
R <- diag(1, I, P)
C <- diag(1, J, Q)
R_inv <- pracma::pinv(t(R) %*% R)
C_inv <- pracma::pinv(t(C) %*% C)
G <- R_inv %*% t(R) %*% DC %*% C %*% C_inv
initM <- R %*% G %*% t(C)
initOmega_hat <- c(0.5, rep((1-0.5)/P, P))
```

```
result <- Log_Likelihood_function_E_ReMI(DC, I, J, initM, initOmega_hat)
result
```

Log_Likelihood_function_REMAXINT

Compute Log-Likelihood for REMAXINT method

Description

Computes the squared error log-likelihood for a given data matrix under the REMAXINT clustering framework.. The function evaluates the fit of the model by comparing the observed data to the reconstructed (fitted) matrix.

Usage

```
Log_Likelihood_function_REMAXINT(DC, I, J, initM)
```

Arguments

DC	A numeric data matrix of size $I \times J$, typically doubly centered, representing the two-mode data to be analyzed.
I	An integer specifying the number of rows (first mode entities) in the data matrix (Dc).
J	An integer specifying the number of columns (second mode entities) in the data matrix (DC).
initM	The matrix M is reconstructed obtained via initial random row partition matrix (R), initial random column partition matrix (C) and initial bi-cluster interaction matrix (G) and it is updated after each updation of R, C and G under the REMAXINT model.

Details

This function is used within the REMAXINT framework to compute the log-likelihood of a model by comparing the original data matrix (DC) to the estimated matrix (*initM*). The error matrix is calculated as the element-wise difference between these two matrices, and the log-likelihood is computed as the sum of squared errors.

Value

A single numeric value representing the log-likelihood (sum of squared residuals).

Note

This function assumes that both input matrices DC and *initM* are of the same dimensions and numeric type.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[REMAXINT](#), [Update_row_clusters_REMAXINT](#), [Update_column_clusters](#), [matrix](#), [sum](#), [pinv](#)

Examples

```
I <- 10; P <- 4; J <- 5; Q <- 2
DC <- matrix(rnorm(I*J), I, J)
R <- diag(1, I, P)
C <- diag(1, J, Q)
R_inv <- pracma::pinv(t(R) %*% R)
C_inv <- pracma::pinv(t(C) %*% C)
G <- R_inv %*% t(R) %*% DC %*% C %*% C_inv
initM <- R %*% G %*% t(C)
result <- Log_Likelihood_function_REMAXINT(DC, I, J, initM)
result
```

Log_LR_Test_Statistic *Compute Log-Likelihood Ratio test statistics for permutation based Extended Random Effect Maximal Interaction (E-ReMI) Two-Mode Clustering Model*

Description

Computes the log-likelihood ratio statistic used for hypothesis testing in the E-ReMI for Two-Mode Clustering model. The statistic compares the fitted model (some intercation) against a null model (no intercation).

Usage

```
Log_LR_Test_Statistic(DC, I, J, initM, initOmega_hat)
```

Arguments

DC	A numeric data matrix of size $I \times J$, typically doubly centered, representing the two-mode data to be analyzed.
I	An integer specifying the number of rows (first mode entities) in the data matrix.
J	An integer specifying the number of columns (second mode entities) in the data matrix.
initM	The matrix M is reconstructed obtained via initial random row partition matrix (R), initial random column partition matrix (C) and initial bi-cluster interaction matrix (G) and it is updated after each updation of R, C and G.
initOmega_hat	A numeric vector of estimated row cluster weights (proportions), summing to 1.

Details

This function evaluates the log-likelihood ratio (LLR) test statistic by comparing the log-likelihood under the fitted model (presence of row by column cluster interaction) to that of a baseline model with no interaction. The test statistic can be used in permutation-based inference procedures to assess significance of row by column cluster interaction in two-mode data.

Value

Returns a single numeric value:

Log_LR The value of the log-likelihood ratio test statistic.

Note

Used internally in permutation-based hypothesis testing procedures for the E-ReMI Two-Mode Clustering framework.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[Permutation_Function](#), [diag](#), [pinv](#)

Examples

```
I <- 10; P <- 4; J <- 5; Q <- 2
DC <- matrix(rnorm(I*J), I, J)
R <- diag(1, I, P)
C <- diag(1, J, Q)
R_inv <- pracma::pinv(t(R) %*% R)
C_inv <- pracma::pinv(t(C) %*% C)
G <- R_inv %*% t(R) %*% DC %*% C %*% C_inv
initM <- R %*% G %*% t(C)
initOmega_hat <- c(0.5, rep((1-0.5)/P, P))
result <- Log_LR_Test_Statistic(DC, I, J, initM, initOmega_hat)
result
```

MaxInt.heatmap.plot *Heatmap for Visualizing Row by Column (Participant-by-Task) Clusters Interaction*

Description

This function generates a heatmap to visualize the estimated interaction values (Γ) between row clusters (e.g., participant groups) and column clusters (e.g., task groups) obtained from the maximal interaction two-mode clustering approach (REMAXINT and/or E-ReMI).

Usage

```
MaxInt.heatmap.plot(Gamma_matrix, row_counts = NULL)
```

Arguments

<code>Gamma_matrix</code>	A numeric matrix containing the estimated interaction values (Γ) for each combination of row and column clusters. Rows represent participant clusters and columns represent task clusters.
<code>row_counts</code>	An optional numeric vector indicating the number of participants in each row cluster. If provided, these counts are displayed in the plot caption.

Details

The function converts the `Gamma_matrix` into a long format and visualizes it as a heatmap using `ggplot2`. The heatmap tiles are colored based on the magnitude of the estimated Γ values, and the corresponding values are displayed within the tiles.

If `row_counts` is provided, the counts for each row cluster are displayed in the plot caption to provide additional context on cluster sizes.

Value

Returns a ggplot object containing the heatmap visualization. The plot includes:

- Heatmap tiles filled based on the Γ values.
- Numeric labels inside each tile showing the rounded Γ value.
- Optional caption with participant counts for each row cluster.

Note

This function is intended to be used after obtaining Γ matrices from the MaxInt_data_analysis function.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[MaxInt_data_analysis](#), [MaxInteraction.plot](#), [screepplot.MaxInt_data_analysis](#)

Examples

```
Gamma_matrix <- matrix(c(0.3, -0.2, 0.5, 0.1, 0.4, -0.1), nrow = 2, byrow = TRUE)
rownames(Gamma_matrix) <- c("Cluster 1", "Cluster 2")
colnames(Gamma_matrix) <- c("Task A", "Task B", "Task C")
MaxInt.heatmap.plot(Gamma_matrix, row_counts = c(50, 40))
```

MaxInteraction.plot *Interaction Plot to understand the nature of row-by-column clusters interaction against the optimal choice of (P) and (Q).*

Description

Generates an interaction plot based on a given interaction matrix (`Gamma_matrix`) resulting from a two-mode clustering analysis. The plot displays estimated interaction effects for the optimal combination of row and column clusters.

Usage

```
MaxInteraction.plot(Gamma_matrix, row_counts = NULL)
```

Arguments

`Gamma_matrix` A numeric matrix containing estimated interaction effects (Γ) between row clusters (e.g., participants) and column clusters (e.g., tasks or items). Rows correspond to row clusters, and columns correspond to column clusters.

`row_counts` An optional numeric vector containing the number of observations (e.g., participants) in each row cluster. If provided, this information is added as a caption in the plot.

Details

The function creates a line plot for each row cluster, where the x-axis corresponds to column clusters and the y-axis to estimated interaction values (Γ). The function automatically assigns generic names (Row Cluster 1, Row Cluster 2, ...) and (Column Cluster 1, Column Cluster 2, ...) if row or column names are not provided.

If `row_counts` is supplied, a caption is added to indicate the number of observations per row cluster.

Value

A ggplot object displaying the interaction plot. The function does not return any additional values.

Note

This function is intended for visualizing the interaction matrix obtained after selecting the optimal number of row and column clusters (P , Q) using `MaxInt_data_analysis` and `MaxInt.Screepplot`.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[MaxInt_data_analysis](#), [screepLOT.MaxInt_data_analysis](#)

Examples

```
Gamma_matrix <- matrix(c(0.5, 0.8, -0.2,
                        0.1, 0.3, 0.6,
                        -0.4, 0.2, 0.7),
                      nrow = 3, byrow = TRUE)
rownames(Gamma_matrix) <- c("High", "Low", "Intermediate")
colnames(Gamma_matrix) <- c("Task 1", "Task 2", "Task 3")
MaxInteraction.plot(Gamma_matrix, row_counts = c(50, 40, 60))
```

MaxInt_data_analysis *Permutation-Based Inference and Interaction Estimation for Maximal Interaction Two-Mode Clustering*

Description

Performs permutation-based statistical inference for the REMAXINT and E-ReMI biclustering models across multiple combinations of row and column clusters. It computes test statistics, p-values, critical values, and estimates interaction parameters for each specified combination of clusters. Furthermore, it provides us optimal row and column partitioning structure for REMAXINT and E-ReMI biclustering methods.

Usage

```
MaxInt_data_analysis(Data, Clus, Nruns, permutations, alpha_level, verbose = TRUE)
```

Arguments

Data	A numeric matrix of dimension $I \times J$, representing the data to be analyzed, where rows correspond to observations (e.g., participants) and columns to variables (e.g., tasks or conditions).
Clus	A matrix with two columns specifying the combinations of row clusters (P) and column clusters (Q) to be analyzed. Each row represents one (P, Q) combination.

Nruns	An integer specifying the number of independent runs for the REMAXINT and E-ReMI algorithms for each (P, Q) combination.
permutations	An integer specifying the number of permutations used to compute the empirical null distribution for test statistics.
alpha_level	Significance level (e.g., 0.05) for computing critical values from the permutation distribution.
verbose	Logical. If TRUE (default), progress messages are printed to the console during execution. Set to FALSE to suppress messages. Messages can also be suppressed using <code>suppressMessages()</code> .

Details

The function iterates over all specified (P, Q) cluster combinations, applies the `Permutation_Function` for each combination, and stores the resulting test statistics, p-values, critical values, and estimated interaction matrices. It further computes a *Gamma value* for each (P, Q) pair, which quantifies the overall interaction strength using row and column cluster cardinalities and the squared interaction parameters within each bicluster.

Value

A list of class "MaxInt_data_analysis" with the following components:

<code>inference_results</code>	A data frame containing test statistics, permutation-based p-values, and critical values for REMAXINT and EReMI.
<code>Gamma_REMAXINT</code>	A list of interaction matrices (Gamma) obtained under REMAXINT for each (P, Q) pair.
<code>Z_REMAXINT</code>	A list of binary row-cluster assignment matrices for REMAXINT.
<code>K_REMAXINT</code>	A list of binary column-cluster assignment matrices for REMAXINT.
<code>Gamma_EReMI</code>	A list of interaction matrices (Gamma) obtained under EReMI for each (P, Q) pair.
<code>Z_EReMI</code>	A list of binary row-cluster assignment matrices for EReMI.
<code>K_EReMI</code>	A list of binary column-cluster assignment matrices for EReMI.
<code>Gamma_results</code>	A data frame with computed interaction Explained Sum of Squares (iESS) for each cluster configuration.

Note

This function serves as a wrapper for repeated calls to `Permutation_Function` to facilitate data analysis for various (P, Q) combinations. This function is computationally intensive and may take time depending on the number of cluster configurations, permutations, and runs specified. Use parallel computing where possible to speed up execution.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[Permutation_Function](#), [screepLOT.MaxInt_data_analysis](#), [MaxInteraction.plot](#)

Examples

```
set.seed(123)
Data <- matrix(rnorm(60), nrow = 10, ncol = 6)
Clus <- matrix(c(2,2, 2,3, 3,2), ncol = 2, byrow = TRUE)

# Run with default verbose output
result <- MaxInt_data_analysis(Data, Clus, Nruns = 5, permutations = 10,
alpha_level = 0.05, verbose = FALSE)
result
```

Permutation_Function *Permutation-based hypothesis testing procedure for REMAXINT and E-REMI based test statistic(s).*

Description

Conducts a permutation test to assess the significance of a latent interaction structure in a data matrix using two model-based clustering approaches: REMAXINT and E-ReMI. It compares the observed likelihood ratio statistics to the distribution under the null hypothesis derived from permutations.

Usage

```
Permutation_Function(D, P, Q, Nruns, permutations, alpha_level)
```


Arguments

D	A numeric matrix of dimension $I \times J$, representing the data to be analyzed.
P	An integer specifying the number of row clusters.
Q	An integer specifying the number of column clusters.
Nruns	The number of REMAXINT and E-REMI algorithms runs to perform for each clustering to avoid local optima.
permutations	The number of permutations used to compute the empirical null distribution.
alpha_level	The significance level (e.g., 0.05) used to determine critical values from the permutation distribution.

Details

This function applies a model-based biclustering algorithm (REMAXINT and E-ReMI) to the original and permuted data matrices to test the presence of an underlying block mean (interaction) structure. For each method, it computes the observed log-likelihood ratio (LLR) statistic and compares it to a distribution of LLRs under the null hypothesis generated through permutation. The procedure yields empirical $\backslash(p \backslash)$ -values and critical values, allowing for hypothesis testing on whether the data contains a meaningful block structure or not.

Value

A list containing the following components:

Obs_Log_LR_REMAXINT	Observed log-likelihood ratio statistic for the REMAXINT method.
Obs_Log_LR_EReMI	Observed log-likelihood ratio statistic for the E-ReMI method.
Perm_Log_LR_REMAXINT	Vector of log-likelihood ratio statistics from permutations for REMAXINT.
Perm_Log_LR_EReMI	Vector of log-likelihood ratio statistics from permutations for E-ReMI.
Crit_Value_Perm_REMAXINT	Critical value at the specified alpha_level from the REMAXINT permutation distribution.
Crit_Value_Perm_EReMI	Critical value at the specified alpha_level from the E-ReMI permutation distribution.
P_value_Perm_REMAXINT	Empirical $\backslash(p \backslash)$ -value from permutation test for REMAXINT.
P_value_Perm_EReMI	Empirical $\backslash(p \backslash)$ -value from permutation test for E-ReMI.
Z_REMAXINT	Maximal row cluster assignment matrix from REMAXINT on observed data.
K_REMAXINT	Maximal column cluster assignment matrix from REMAXINT on observed data.
Z_EReMI	Maximal row cluster assignment matrix from E-ReMI on observed data.
K_EReMI	Maximal column cluster assignment matrix from E-ReMI on observed data.
Gamma_REMAXINT	Estimated block mean (bi-cluster interaction) matrix from REMAXINT.
Gamma_EReMI	Estimated block mean (bi-cluster interaction) matrix from E-ReMI.

Note

Permutation function ensures input matrix D is preprocessed appropriately (e.g., doubly centered) within this function.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[REMAXINT](#), [E_ReMI](#), [Log_LR_Test_Statistic](#)

Examples

```
I <- 10; P <- 4; J <- 5; Q <- 2; Nruns <- 5; permutations <- 10; alpha_level <- 0.05
D <- matrix(rnorm(I*J), I, J)
result <- Permutation_Function(D, P, Q, Nruns, permutations, alpha_level)
result
```

```
print.MaxInt_data_analysis
```

Print Summary for MaxInt Data Analysis Results

Description

Displays a concise summary of the results from a `MaxInt_data_analysis` object, including inference statistics and interaction Explained Sum of Squares (iESS).

Usage

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

Arguments

- x An object of class "MaxInt_data_analysis", typically returned by the [MaxInt_data_analysis](#) function.
- ... Additional arguments passed to or from other methods (currently ignored).

Details

This method provides a simple print interface for viewing the key components of a MaxInt_data_analysis object. It prints the results of the permutation-based inference tests and the computed Gamma_results, which contain the interaction Explained Sum of Squares (iESS) for each cluster combinations.

Value

Returns invisible(x). Called primarily for its side effects (printing to the console).

Note

This is the default print method for objects of class "MaxInt_data_analysis". Use it to quickly inspect inference and iESS results.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation-based alternative. *Available on OSF* (submitted to ADAC). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, 40, 298-331. Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, 15(4), 987-1013.

See Also

[MaxInt_data_analysis](#)

Examples

```
set.seed(123)
Data <- matrix(rnorm(60), nrow = 10, ncol = 6)
Clus <- matrix(c(2,2, 2,3, 3,2), ncol = 2, byrow = TRUE)
result <- MaxInt_data_analysis(Data, Clus, Nruns = 5, permutations = 10, alpha_level = 0.05)
result
```

Randpartition_function

Generate a Random Partition Matrix of Equally sized row and column clusters. These clusters are mutually exclusive and there is no empty row and column cluster.

Description

Generates a binary matrix representing a random partition of i items into p latent groups. Each row indicates group membership in a one-hot encoded format.

Usage

```
Randpartition_function(i, p)
```

Arguments

i	An integer specifying the total number of rows or columns of a data matrix to be partitioned.
p	An integer specifying the number of latent row or column clusters.

Details

This function creates a binary matrix of size $i \times p$, where each row corresponds to a row cluster. The function first ensures that each cluster has at least one initial member (by using an identity matrix), and then assigns the remaining items randomly to clusters. Finally, the matrix rows are shuffled to ensure randomization of the assignment order.

Value

A binary matrix of dimensions $i \times p$:

A	A binary matrix indicating the membership of each row to one of the p row clusters. Each row has exactly one 1 indicating cluster membership.
-----	---

Note

This function is primarily used for generating initial row or column cluster membership matrices in the REMAXINT and E-ReMI methods.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[REMAXINT](#), [E_ReMI](#), [diag](#)

Examples

```
result <- Randompartition_function(i = 10, p = 3)
result
```

REMAXINT

Two-Mode Clustering via the REMAXINT Method

Description

Performs two-mode clustering (row and column) of a data matrix using the REMAXINT (Random Effect Maximal Interaction Two-mode Clustering) algorithm. It iteratively updates row and column cluster memberships to maximize a Gaussian log-likelihood criterion and ultimately provides partitions of the data that maximize interaction between row and column clusters. Furthermore, REMAXINT method assumes row and column clusters sizes are equal in expectation.

Usage

```
REMAXINT(DC, P, Q, Nruns)
```

Arguments

DC	A doubly centered numeric matrix of dimension $I \times J$ representing the data to be clustered.
P	An integer indicating the number of row clusters.
Q	An integer indicating the number of column clusters.
Nruns	An integer specifying the number of random initializations (runs) to perform. The best solution (based on log-likelihood) is returned.

Details

The REMAXINT function implements a likelihood-based biclustering approach. For each of the specified N_{runs} , it initializes row and column cluster assignments randomly and iteratively updates them to maximize the log-likelihood of the data under a block interaction model. The algorithm uses alternating updates of the row and column cluster assignments, along with an estimation of the latent cluster interaction matrix G and noise variance. To avoid poor local maxima, multiple random starts are performed, and the result with the highest final log-likelihood is selected.

Value

A list with the following components:

BestR	An $I \times P$ binary matrix indicating maximal row cluster assignments.
BestC	A $J \times Q$ binary matrix indicating maximal column cluster assignments.
BestG	A $P \times Q$ matrix representing estimated latent interactions between row and column clusters.
LL	The maximized log-likelihood value corresponding to the maximal clustering solution.

Note

This method is particularly useful when the data shows block-wise latent interaction structures between row and column clusters.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[Permutation_Function](#), [Randompartition_function](#), [Update_row_clusters_REMAXINT](#), [Update_column_clusters](#), [Update_G_Omega](#), [Log_Likelihood_function_REMAXINT](#),

Examples

```
I <- 10; P <- 4; J <- 5; Q <- 2; Nruns <- 5
DC <- matrix(rnorm(I*J), I, J)
result <- REMAXINT(DC, P, Q, Nruns)
result
```

```
screepLOT.MaxInt_data_analysis
```

Scree Plot for Maximal Interaction Two-Mode Clustering Results

Description

Generates a scree plot showing the interaction Explained Sum of Squares (iESS) for different combinations of latent row and column clusters in the Maximal Interaction two-mode clustering analysis.

Usage

```
## S3 method for class 'MaxInt_data_analysis'
screepLOT(x, ...)
```

Arguments

x An object of class `MaxInt_data_analysis`, typically returned by the function that performs the Maximal Interaction two-mode clustering. It should contain a list element `Gamma_results` with rows `P`, columns `Q`, and cell entries `Gamma_value`.

... Additional graphical parameters passed to the underlying plot function.

Details

This function helps visualize how the interaction Explained Sum of Squares (iESS) varies across different combinations of latent row and column cluster numbers, `P` and `Q`, by plotting their product $(P - 1)(Q - 1)$ on the x-axis and the corresponding iESS on the y-axis. This plot aids in selecting the optimal number of latent row by column clusters.

Value

Returns a base R plot object for visual inspection. No return value is assigned invisibly.

If needed, additional customization can be applied using graphical parameters through the `...` argument.

Note

Ensure that the input object contains a properly structured `Gamma_results` data frame. Use this plot to inform your choice of latent clusters.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to Advances in Data Analysis and Classification). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[MaxInt_data_analysis](#), [MaxInteraction.plot](#)

Examples

```
set.seed(123)
set.seed(123)
Data <- matrix(rnorm(60), nrow = 10, ncol = 6)
Clus <- matrix(c(2,2, 2,3, 3,2), ncol = 2, byrow = TRUE)
result <- MaxInt_data_analysis(Data, Clus, Nruns = 5, permutations = 10, alpha_level = 0.05)
screplot(result)
```

Unequal_Randompartition_Function

Generate a Random Partition Matrix of Unequally sized row clusters. Therefore, it assign higher weight or probability to first the row cluster and the remaining row clusters are evenly distributed.

Description

Generates a binary matrix representing a random partition of i rows into p row groups, ensuring each group is represented at least once, and allowing the first group to have a higher probability.

Usage

```
Unequal_Randompartition_Function(i, p, firstp)
```


Arguments

<code>i</code>	An integer specifying the total number of items (rows) to partition.
<code>p</code>	An integer specifying the number of clusters or groups.
<code>firstp</code>	A numeric value (between 0 and 1) indicating the selection probability for the first group. The remaining probability is evenly distributed among the remaining $p-1$ groups.

Details

The function first initializes a $p \times p$ identity matrix to ensure each group is selected at least once. The rest of the rows ($i - p$) are assigned to groups based on a multinomial distribution with a specified probability vector, where the first group's probability is `firstp` and the rest share the remaining $1 - \text{firstp}$ equally.

Value

A binary matrix of dimension $i \times p$, where each row contains a single 1 indicating the assigned group, and 0s elsewhere.

Note

This function is primarily used for generating initial row cluster membership matrix in the E-ReMI methods.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[E_ReMI](#), [rmultinom](#), [diag](#), [sample](#)

Examples

```
result <- Unequal_Randompartition_Function(i = 20, p = 4, firstp = 0.5)
result
```

Update_column_clusters

Update Column Cluster Assignments in REMAXINT and E-ReMI clustering framework.

Description

Updates column cluster assignments (C matrix) and corresponding G matrix in a two-mode clustering model based on a likelihood maximization procedure.

Usage

```
Update_column_clusters(DC, I, J, initR, initC, initG, Q)
```

Arguments

DC	A numeric data matrix of size $I \times J$, typically doubly centered, representing the two-mode data to be analyzed.
I	The number of rows (first mode entities) in the data matrix.
J	The number of columns (second mode entities) in the data matrix.
initR	Initial binary row-cluster assignment matrix of size $I \times P$.
initC	Initial binary column-cluster assignment matrix of size $J \times Q$.
initG	Initial cluster interaction matrix of size $P \times Q$.
Q	The number of column clusters.

Details

This function updates the column cluster assignments by iteratively testing all possible cluster configurations for each column and selecting the one that maximizes the log-likelihood. After reassigning the clusters, it checks for empty clusters and reassigns the least likely columns to ensure all clusters are non-empty. The function then updates the G matrix accordingly.

Value

A list with the following components:

C	The updated column cluster assignment matrix of size $J \times Q$.
G	The updated cluster interaction matrix of size $P \times Q$.

Note

This function assumes that the input cluster assignment matrices are binary and mutually exclusive for each item (i.e., hard clustering).

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[REMAXINT](#), [E_ReMI](#), [pinv](#)

Examples

```
I <- 10
P <- 4
J <- 5
Q <- 2
Nruns <- 5
DC <- matrix(rnorm(I * J), I, J)
initR <- diag(1, nrow = I, ncol = P)
initC <- diag(1, nrow = J, ncol = Q)
R_inv <- pracma::pinv(t(initR) %*% initR)
C_inv <- pracma::pinv(t(initC) %*% initC)
initG <- R_inv %*% t(initR) %*% DC %*% initC %*% C_inv
initG <- as.matrix(initG)
result <- Update_column_clusters(DC, I, J, initR, initC, initG, Q)
result
```

Update_G_Omega

Updates Row Cluster Membership Probabilities

Description

Updates the Row Cluster Membership Probabilities based on the current assignments of rows to row clusters in E_ReMI clustering algorithm.

Usage

Update_G_Omega(DC, I, J, Updated_C, Updated_R)

Arguments

DC	A numeric data matrix of size $I \times J$, typically doubly centered, representing the two-mode data to be analyzed.
I	An integer specifying the number of row objects in the data matrix.
J	An integer specifying the number of column variables in the data matrix.
Updated_C	A binary matrix indicating the current cluster membership of columns. Each row corresponds to a column in DC, and each column represents a column cluster.
Updated_R	A binary matrix indicating the current cluster membership of rows. Each row corresponds to a row in DC, and each column represents a row cluster.

Details

This function is typically called within each iteration of a E-ReMI clustering algorithm, where the goal is to maximize the log-likelihood or minimize least squared of DC via row and column cluster memberships. Thus it updates model parameters (row cluster membership probabilities) under E-ReMI clustering framework.

Value

A list with the following components:

Omega	A numeric vector of updated row cluster membership probabilities.
G	The updated row by column clusters interaction matrix.
M	The reconstructed data matrix from the current estimates.
Sigma	The updated estimate of residual variance.

Note

This function assumes binary (one-hot encoded) cluster membership matrices Updated_C and Updated_R.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[E_ReMI](#), [REMAXINT](#), [pinv](#)

Examples

```
I <- 10; P <- 4; J <- 5; Q <- 2
DC <- matrix(rnorm(I*J), I, J)
Updated_R <- diag(1, I, P)
Updated_C <- diag(1, J, Q)
result <- Update_G_Omega(DC, I, J, Updated_C, Updated_R)
result
```

Update_row_clusters_E_ReMI

Update Row Cluster Assignments in E-ReMI clustering framework.

Description

Updates the row cluster membership matrix within the E-ReMI algorithm by maximizing the log-likelihood for each row based on the current estimates of model parameters.

Usage

```
Update_row_clusters_E_ReMI(DC, I, J, P,
Updated_R, Updated_C, Updated_omegahat,
Updated_G,
Updated_sigma)
```

Arguments

DC	A doubly centered numeric matrix of dimension $I \times J$ representing the data matrix to be clustered.
I	An integer specifying the number of rows (objects) in the data matrix.
J	An integer specifying the number of columns (features) in the data matrix.
P	An integer specifying the number of row clusters.
Updated_R	A binary $I \times P$ matrix representing the current row cluster membership matrix.
Updated_C	A binary $J \times Q$ matrix representing the current column cluster membership matrix.
Updated_omegahat	A numeric vector of length P representing the current estimated mixing proportions for row clusters.
Updated_G	A numeric matrix of dimension $P \times Q$ representing the current latent structure or interaction pattern between row and column clusters.
Updated_sigma	A numeric value representing the current estimate of error variance or noise parameter in the model.

Details

This function iteratively updates the row cluster assignments for each row in the data matrix DC. For each row, it evaluates the log-likelihood of assignment to all possible clusters and updates the cluster membership to the one that maximizes the likelihood. Singleton constraints are respected to avoid assigning rows to clusters that would result in empty or singleton groups.

Value

A binary matrix of the same dimensions as Updated_R, representing the updated row cluster membership after one iteration of the E-ReMI update step.

Note

This function is a part of the E-ReMI clustering framework.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[E_ReMI](#), [pinv](#)

Examples

```
I <- 10; P <- 4; J <- 5; Q <- 2
DC <- matrix(rnorm(I*J), I, J)
Updated_R <- diag(1, I, P)
Updated_C <- diag(1, J, Q)
Updated_omegahat <- c(0.5, rep((1-0.5)/P, P))
R_pinv <- pracma::pinv(t(Updated_R) %*% Updated_R)
C_pinv <- pracma::pinv(t(Updated_C) %*% Updated_C)
Updated_G <- R_pinv %*% t(Updated_R) %*% DC %*% Updated_C %*% C_pinv
Updated_G <- as.matrix(Updated_G)
Updated_M <- Updated_R %*% Updated_G %*% t(Updated_C)
Updated_sigma <- sum((DC - Updated_M)^2)/(I*J)
```

```

result <- Update_row_clusters_E_ReMI(
  DC, I, J, P,
  Updated_R, Updated_C,
  Updated_omegahat, Updated_G,
  Updated_sigma
)
result

```

Update_row_clusters_REMAXINT

Update Row Cluster Assignments in REMAXINT clustering method

Description

Updates the row cluster assignments for a fixed column clustering and G matrix in the REMAXINT clustering framework, using a maximum likelihood strategy.

Usage

```
Update_row_clusters_REMAXINT(DC, I, J, P, Updated_R, Updated_C, Updated_G)
```

Arguments

DC	The data matrix of dimensions $I \times J$, representing the doubly centering input data to be clustered.
I	The number of rows (objects) in the data matrix (DC).
J	The number of columns (features/conditions) in the data matrix (DC).
P	The number of row clusters.
Updated_R	The current binary row cluster membership matrix of size $I \times P$, where each row has a 1 in the assigned cluster (column).
Updated_C	The fixed binary column cluster membership matrix of size $J \times Q$.
Updated_G	The matrix representing latent interactions between row and column clusters of size $P \times Q$.

Details

The function iteratively evaluates each row in the data matrix and reassigns it to the cluster that maximizes the log-likelihood under the REMAXINT model. Rows assigned to singleton clusters are not re-evaluated. The function avoids assigning rows to clusters with only one member (singleton clusters) to prevent degenerate solutions.

Value

Returns an updated row cluster membership matrix of the same dimension as Updated_R, where each row is reassigned based on the maximum likelihood estimate given the current column clustering and latent interactions between row and column clusters of size $P \times Q$. This matrix has:

Updated_R	The updated binary row cluster membership matrix.
-----------	---

Note

Used internally by the REMAXINT model fitting procedure. Assumes fixed column cluster membership and updates row assignments to maximize model fit.

Author(s)

Zaheer Ahmed (ahmed@ifado.de), Jan Schepers (jan.schepers@maastrichtuniversity.nl), Uwe Ligges (ligges@statistik.tu-dortmund.de), Alberto Cassese (alberto.cassese@unifi.it), Gerard van Breukelen (gerard.vbreukelen@maastrichtuniversity.nl), Katja Ickstadt (ickstadt@statistik.tu-dortmund.de), Edmund Wascher (wascher@ifado.de)

References

Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2025). Robustness study of normality-based likelihood ratio tests for testing maximal interaction two-mode clustering and a permutation based alternative. *Available on OSF* (submitted to *Advances in Data Analysis and Classification*). Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2023). E-ReMI: extended maximal interaction two-mode clustering. *Journal of Classification*, *40*, 298-331 Ahmed, Z., van Breukelen, G. J. P., Schepers, J., & Cassese, A. (2021). REMAXINT: a two-mode clustering-based method for statistical inference on two-way interaction. *Advances in Data Analysis and Classification*, *15*(4), 987-1013

See Also

[REMAXINT, pinv](#)

Examples

```
I <- 10; P <- 4; J <- 5; Q <- 2
DC <- matrix(rnorm(I*J), I, J)
Updated_R <- diag(1, I, P)
Updated_C <- diag(1, J, Q)
R_pinv <- pracma::pinv(t(Updated_R) %** Updated_R)
C_pinv <- pracma::pinv(t(Updated_C) %** Updated_C)
Updated_G <- R_pinv %** t(Updated_R) %** DC %** Updated_C %** C_pinv
Updated_G <- as.matrix(Updated_G)
result <- Update_row_clusters_REMAXINT(DC, I, J, P, Updated_R, Updated_C, Updated_G)
result
```


Index

- * **E-ReMI based two-mode clustering**
 - Log_LR_Test_Statistic, 9
 - * **E-ReMI**
 - E_ReMI, 4
 - Log_Likelihood_function_E_ReMI, 6
 - * **Maximal interaction**
 - E_ReMI, 4
 - MaxInt_data_analysis, 14
 - MaxIntTools-package, 2
 - REMAXINT, 21
 - * **Permutation based hypothesis testing**
 - Log_LR_Test_Statistic, 9
 - * **Permutation based inference**
 - MaxInt_data_analysis, 14
 - MaxInteraction.plot, 13
 - * **REMAXINT and E-ReMI based two-mode clustering**
 - Randompartition_function, 20
 - * **REMAXINT and E-ReMI**
 - MaxInt.heatmap.plot, 11
 - MaxInt_data_analysis, 14
 - MaxInteraction.plot, 13
 - * **REMAXINT**
 - Log_Likelihood_function_REMAXINT, 8
 - REMAXINT, 21
 - * **Two-mode clustering**
 - E_ReMI, 4
 - * **Two-mode data**
 - REMAXINT, 21
 - * **bi-clustering**
 - MaxInt.heatmap.plot, 11
 - * **biclustering**
 - MaxInt_data_analysis, 14
 - * **column clustering in REMAXINT and E-REMI**
 - Update_column_clusters, 26
 - * **likelihood**
 - Log_Likelihood_function_E_ReMI, 6
 - Log_Likelihood_function_REMAXINT, 8
 - * **methods**
 - print.MaxInt_data_analysis, 18
 - * **optimal clustering**
 - screepplot.MaxInt_data_analysis, 23
 - * **print**
 - print.MaxInt_data_analysis, 18
 - * **random partition matrix**
 - Randompartition_function, 20
 - * **random row partition matrix in E_ReMI**
 - Unequal_Randompartition_Function, 24
 - * **row cluster membership matrix**
 - Update_G_Omega, 27
 - * **unequally sized row clusters**
 - Log_Likelihood_function_E_ReMI, 6
 - * **updating row cluster membership in E-ReMI**
 - Update_row_clusters_E_ReMI, 29
 - Update_row_clusters_REMAXINT, 31
 - * **visualization optimal interaction**
 - MaxInteraction.plot, 13
 - * **visualization**
 - MaxInt.heatmap.plot, 11
 - screepplot.MaxInt_data_analysis, 23
- diag, 11, 21, 25
- E_ReMI, 4, 4, 7, 18, 21, 25, 27, 29, 30
- extract_Gamma_EReMI, 3
- Log_Likelihood_function_E_ReMI, 6, 6
- Log_Likelihood_function_REMAXINT, 8, 22
- Log_LR_Test_Statistic, 9, 18
- matrix, 9
- MaxInt.heatmap.plot, 3, 11
- MaxInt_data_analysis, 3, 4, 12, 14, 14, 19, 24

MaxInteraction.plot, [3](#), [4](#), [12](#), [13](#), [16](#), [24](#)
MaxIntTools-package, [2](#)

Permutation_Function, [4](#), [6](#), [11](#), [16](#), [16](#), [22](#)
pinv, [6](#), [7](#), [9](#), [11](#), [27](#), [29](#), [30](#), [32](#)
print.MaxInt_data_analysis, [18](#)

Randompartition_function, [6](#), [20](#), [22](#)
REMAXINT, [9](#), [18](#), [21](#), [21](#), [27](#), [29](#), [32](#)
rmultinom, [25](#)

sample, [25](#)
screepLOT.MaxInt_data_analysis, [3](#), [4](#), [12](#),
[14](#), [16](#), [23](#)
sum, [9](#)

Unequal_Randompartition_Function, [6](#), [24](#)
Update_column_clusters, [6](#), [9](#), [22](#), [26](#)
Update_G_Omega, [6](#), [22](#), [27](#)
Update_row_clusters_E_ReMI, [6](#), [7](#), [29](#)
Update_row_clusters_REMAXINT, [9](#), [22](#), [31](#)