Package: mimdo (via r-universe)

August 22, 2024

Type Package

Title Multivariate Imputation by 'Mahalanobis' Distance Optimization

Version 0.1.0

Author Geovert John Labita [aut, cre]

Maintainer Geovert John Labita <geovertjohn.labita@ustp.edu.ph>

Description Imputes missing values of an incomplete data matrix by minimizing the 'Mahalanobis' distance of each sample from the overall mean.

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.1

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

RemoteUrl https://github.com/gjlabita/mimdo

RemoteRef HEAD

RemoteSha c1c331cfc686719868b053cd2436edb152f23dd7

Contents

`mimdo` *Multivariate Imputation by Mahalanobis Distance Optimization*

Description

Imputes missing values of an incomplete data matrix by minimizing the Mahalanobis distance of each sample from the overall mean. By utilizing Mahalanobis distance, this imputation method is preferable to be used on datasets with highly correlated variables.

Usage

```
mimdo(incomplete_data, inverse, iterations = 30)
```

Arguments

`incomplete_data`
: A data frame with missing values.

`inverse`
: If TRUE, the inverse covariance matrix will be used for distance calculation. If the covariance matrix is non-invertible, use inverse = FALSE.

`iterations`
: Number of iterations. It can be adjusted to avoid long running time.

Details

The output is a complete imputed data matrix.

Author(s)

Geovert John D. Labita

References

Labita, GJ.D. and Tubo, B.F. (2024). Missing data imputation via optimization approach: An application to K-means clustering of extreme temperature. Reliability: Theory and Applications, 2(78), 115-123. DOI: https://doi.org/10.24412/1932-2321-2024-278-115-123

Bertsimas, D., Pawlowski, C., and Zhou, Y.D. (2018). From predictive methods to missing data imputation: An optimization approach. Journal of Machine Learning Research, 18(196), 1-39.

Examples

```
incomplete_data<-as.data.frame(matrix(c(5.1,NA,4.7,NA,3.0,3.2,1.4,1.4,NA,0.2,0.2,NA),nrow=3))
mimdo(incomplete_data, inverse=FALSE)
```

Index