Package ‘GeneralizedUmatrixGPU’

November 17, 2025

Type Package

Title Credible Visualization for Two-Dimensional Projections of Data

Version 0.1.8

Date 2025-11-04

Maintainer Quirin Stier <Quirin_Stier@gmx.de>

Description Projections are common dimensionality reduction methods, which represent high-dimensional data in a two-dimensional space. However, when restricting the output space to two dimensions, which results in a two dimensional scatter plot (projection) of the data, low dimensional similarities do not represent high dimensional distances coercively [Thrun, 2018] <DOI:10.1007/978-3-658-20540-9>. This could lead to a misleading interpretation of the underlying structures [Thrun, 2018]. By means of the 3D topographic map the generalized Umatrix is able to depict errors of these two-dimensional scatter plots. The package is derived from the book of Thrun, M.C.: ``Projection Based Clustering through Self-Organization and Swarm Intelligence'' (2018) <DOI:10.1007/978-3-658-20540-9> and the main algorithm called simplified self-organizing map for dimensionality reduction methods is published in Thrun, M.C. and Ultsch, A.: ``Uncovering High-dimensional Structures of Projections from Dimensionality Reduction Methods'' (2020) <DOI:10.1016/j.mex.2020.101093>.

License GPL-3

Imports Rcpp (>= 1.0.8), RcppParallel (>= 5.1.4), ggplot2, GeneralizedUmatrix

Suggests DataVisualizations, rgl, grid, mgcv, png, reshape2, fields, ABCanalysis, plotly, deldir, methods, knitr (>= 1.12), rmarkdown (>= 0.9), ProjectionBasedClustering

LinkingTo Rcpp, RcppArmadillo, RcppParallel

Depends R (>= 3.0)

NeedsCompilation yes

SystemRequirements GNU make, pandoc (>=1.12.3, needed for vignettes), OpenCL shared library (provided by an SDK such as AMD/NVIDIA)

LazyLoad yes

LazyData TRUE

Encoding UTF-8

Author Quirin Stier [aut, cre] (ORCID:
<https://orcid.org/0000-0002-7896-4737>),
Michael Thrun [aut, cph] (ORCID:
<https://orcid.org/0000-0001-9542-5543>),
The Khronos Group Inc. [cph]

Repository CRAN

Date/Publication 2025-11-17 20:50:02 UTC

Contents

`GeneralizedUmatrixGPU-package`
Credible Visualization for Two-Dimensional Projections of Data

---

Description

Projections are common dimensionality reduction methods, which represent high-dimensional data in a two-dimensional space. However, when restricting the output space to two dimensions, which results in a two dimensional scatter plot (projection) of the data, low dimensional similarities do not represent high dimensional distances coercively [Thrun, 2018] <DOI: 10.1007/978-3-658-20540-9>. This could lead to a misleading interpretation of the underlying structures [Thrun, 2018]. By means of the 3D topographic map the generalized Umatrix is able to depict errors of these two-dimensional scatter plots. The package is derived from the book of Thrun, M.C.: "Projection Based Clustering through Self-Organization and Swarm Intelligence" (2018) <DOI:10.1007/978-3-658-20540-9> and the main algorithm called simplified self-organizing map for dimensionality reduction methods is published in Thrun, M.C. and Ultsch, A.: "Uncovering High-dimensional Structures of Projections from Dimensionality Reduction Methods" (2020) <DOI:10.1016/j.mex.2020.101093>.

Details

For a brief introduction to **GeneralizedUmatrixGPU** please see the vignette Introduction of the Generalized Umatrix Package.

For further details regarding the generalized Umatrix see [Thrun, 2018], chapter 4-5, or [Thrun/Ultsch, 2020].

If you want to verifiy your clustering result externally, you can use `Heatmap` or `SilhouettePlot` of the CRAN package `DataVisualizations`.

Index of help topics:

```
Chainlink                 Chainlink is part of the Fundamental Clustering
                          Problem Suit (FCPS) [Thrun/Ultsch, 2020].
DefaultColorSequence      Default color sequence for plots
GeneralizedUmatrixGPU     Generalized U-Matrix on GPU for Projection
                          Methods published in [Thrun/Ultsch, 2020]
GeneralizedUmatrixGPU-package
                          Credible Visualization for Two-Dimensional
                          Projections of Data
```

Author(s)

Michal Thrun

Maintainer: Michael Thrun <mthrun@informatik.uni-marburg.de>

References

[Thrun/Ultsch, 2020] Thrun, M. C., & Ultsch, A.: Uncovering High-Dimensional Structures of Projections from Dimensionality Reduction Methods, MethodsX, Vol. 7, pp. 101093, DOI doi:10.1016/j.mex.2020.101093, 2020.

[Thrun, 2018] Thrun, M. C.: Projection Based Clustering through Self-Organization and Swarm Intelligence, doctoral dissertation 2017, Springer, Heidelberg, ISBN: 978-3-658-20539-3, doi:10.1007/9783658205409, 2018.

[Ultsch/Thrun, 2017] Ultsch, A., & Thrun, M. C.: Credible Visualizations for Planar Projections, in Cottrell, M. (Ed.), 12th International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM), IEEE Xplore, France, 2017.

Examples

```
library(GeneralizedUmatrix)
data("Chainlink")
Data=Chainlink$Data
Cls=Chainlink$Cls
InputDistances=as.matrix(dist(Data))
res=cmdscale(d=InputDistances, k = 2, eig = TRUE, add = FALSE, x.ret = FALSE)
ProjectedPoints=as.matrix(res$points)
#see also ProjectionBasedClustering package for other common projection methods
#see DatabionicSwarm for projection method without parameters or objective function
# ProjectedPoints=DatabionicSwarm::Pswarm(Data)$ProjectedPoints

resUmatrix=GeneralizedUmatrixGPU(Data,ProjectedPoints)
plotTopographicMap(resUmatrix$Umatrix,resUmatrix$Bestmatches,Cls)
```

`Chainlink` *Chainlink is part of the Fundamental Clustering Problem Suit (FCPS) [Thrun/Ultsch, 2020].*

Description

linear not separable dataset of two interwined chains.

Usage

```
data("Chainlink")
```

Details

Size 1000, Dimensions 3, stored in `Chainlink$Data`

Teo clusters, stored in `Chainlink$Cls`

Published in [Ultsch et al.,1994] in German and [Ultsch 1995] in English.

References

[Thrun/Ultsch, 2020] Thrun, M. C., & Ultsch, A.: Clustering Benchmark Datasets Exploiting the Fundamental Clustering Problems, Data in Brief,Vol. 30(C), pp. 105501, DOI 10.1016/j.dib.2020.105501 , 2020.

[Ultsch 1995] Ultsch, A.: Self organizing neural networks perform different from statistical k-means clustering, Proc. Society for Information and Classification (GFKL), Vol. 1995, Basel 8th-10th March, 1995.

[Ultsch et al.,1994] Ultsch, A., Guimaraes, G., Korus, D., & Li, H.: Knowledge extraction from artificial neural networks and applications, Parallele Datenverarbeitung mit dem Transputer, pp. 148-16Chainlink, Springer, 1994.

Examples

```
data(Chainlink)
str(Chainlink)

## Not run:
require(DataVisualizations)
DataVisualizations::Plot3D(Chainlink$Data,Chainlink$Cls)

## End(Not run)
```

`DefaultColorSequence` *Default color sequence for plots*

Description

Defines the default color sequence for plots made within the Projections package.

Usage

```
data("DefaultColorSequence")
```

Format

A vector with 562 different strings describing colors for plots.

`GeneralizedUmatrixGPU` *Generalized U-Matrix on GPU for Projection Methods published in [Thrun/Ultsch, 2020]*

Description

Generalized U-Matrix visualizes high-dimensional distance and density based structurs in two-dimensional scatter plots of projectios methods like CCA, MDS, PCA or NeRV [Ultsch/Thrun, 2017] with the help of a topographic map with hypsometrioc tints [Thrun et al. 2016] using a simplified emergent SOM published in [Thrun/Ultsch, 2020].

Usage

```
GeneralizedUmatrixGPU(Data, ProjectedPoints, PlotIt = FALSE, Cls = NULL,
Toroid = TRUE, Tiled = FALSE, DataPerEpoch = 1, Verbose = 0, ...)
```

Arguments

`Data` [1:n,1:d] array of data: n cases in rows, d variables in columns.

`ProjectedPoints` [1:n,2] matrix containing coordinates of the Projection: A matrix of the fitted configuration.

`PlotIt` Optional,bool, defaut=FALSE, if =TRUE: U-Marix of every current Position of Databots will be shown.

`Cls` Optional, For plotting, see `plotUmatrix` in package Umatrix.

`Toroid` Optional, Default=TRUE, ==FALSE planar computation with borders defined by projection method ==TRUE: toroid borderless (toroidal) computation, the four borders defined by projection method are ignored.

`Tiled` Optional,For plotting see `plotUmatrix` in package Umatrix

DataPerEpoch	Optional, scalar, value above zero and below 1 starts sampling and defines percentage of data points sampled in each epoch during the learning phase. Beware: Experimental!
Verbose	Integer, determining text output during computation (Verbose > 0) or silent mode (Verbose=0).
...	Further parameters.

Details

Introduced first in the PhD thesis in [Thrun, 2018, p.46]. Furthermore the two parts of the work were peer-reviewed and published in [Ultsch/Thrun, 2017, Thrun/Ultsch, 2020].

Value

List with

Umatrix	[1:Lines,1:Columns] Umatrix to be plotted, numerical matrix storing the U-heights, see [Thrun, 2018] for definition.
EsomNeurons	[1:Lines,1:Columns,1:weights] 3-dimensional numeric array (wide format), not wts (long format).
Bestmatches	[1:n,1:2] Positions of GridConverted Projected Points on the Umatrix to the predefined Grid by Lines and Columns, First Columns has the content of the Line No and second Column of the Column number.
sESOMparamaters	internals for debugging
Lines	Number of Lines
Columns	Number of Columns
gplotres	output of ggplot2

Note

With the update of 01.01.2024, version 1.3 a minor change is included that is not mentioned in the referenced papers: for large number of cases and small radii the learning rate decays to 0.1 instead of remaining constant (any other case).

Author(s)

Quirin Stier, Michael Thrun

References

[Thrun et al., 2016] Thrun, M. C., Lerch, F., Loetsch, J., & Ultsch, A.: Visualization and 3D Printing of Multivariate Data of Biomarkers, in Skala, V. (Ed.), International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG), Vol. 24, Plzen, http://wscg.zcu.cz/wscg2016/short/A43-full.pdf, 2016.

[Thrun, 2018] Thrun, M. C.: Projection Based Clustering through Self-Organization and Swarm Intelligence, doctoral dissertation 2017, Springer, Heidelberg, ISBN: 978-3-658-20539-3, doi:10.1007/9783658205409, 2018.

[Ultsch/Thrun, 2017] Ultsch, A., & Thrun, M. C.: Credible Visualizations for Planar Projections, in Cottrell, M. (Ed.), 12th International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM), IEEE Xplore, France, 2017.

[Thrun/Ultsch, 2020] Thrun, M. C., & Ultsch, A.: Uncovering High-Dimensional Structures of Projections from Dimensionality Reduction Methods, MethodsX, Vol. 7, pp. 101093, DOI doi:10.1016/j.mex.2020.101093, 2020.

Examples

```
library(GeneralizedUmatrix)
data("Chainlink")
Data=Chainlink$Data
Cls=Chainlink$Cls
InputDistances=as.matrix(dist(Data))
res=cmdscale(d=InputDistances, k = 2, eig = TRUE, add = FALSE, x.ret = FALSE)
ProjectedPoints=as.matrix(res$points)

if(requireNamespace("ProjectionBasedClustering")){
Stress = ProjectionBasedClustering::KruskalStress(InputDistances,
as.matrix(dist(ProjectedPoints)))
}


resUmatrix=GeneralizedUmatrixGPU(Data[1:2,], ProjectedPoints[1:2,])
#plotTopographicMap(resUmatrix$Umatrix,resUmatrix$Bestmatches)
#testing takes longer than 5 secs


resUmatrix=GeneralizedUmatrixGPU(Data,ProjectedPoints)
#plotTopographicMap(resUmatrix$Umatrix,resUmatrix$Bestmatches,Cls)
```

Index