

Literature summaries

Adrià Casanova Lloveras

February 10, 2025

Contents

1	Rdimtools: An R Package for Dimension Reduction and Intrinsic Dimension Estimation [1]	2
1.1	Abstract	2
1.2	Key Points	2
1.3	Example R Code	2

1 Rdimtools: An R Package for Dimension Reduction and Intrinsic Dimension Estimation [1]

1.1 Abstract

Discovering patterns of the complex high-dimensional data is one of the fundamental pillars of modern data science. Dimension reduction and intrinsic dimension estimation are two thematic programs that facilitate geometric characterization of the data. We present Rdimtools, an R package that supports 143 dimension reduction and manifold learning methods and 17 dimension estimation algorithms whose unprecedented extent makes multifaceted scrutiny of the data in one place easier. Rdimtools is distributed under the MIT license and is accessible from CRAN, GitHub, and its package website, all of which deliver instruction for installation, self-contained examples, and API documentation.

1.2 Key Points

- **R package:** that supports 143 dimension reduction and manifold learning methods and 17 dimension estimation algorithms.
- **Other libraries:** drtoolbox in MATLAB, scikit-learn in Python, a C++ template library tapkee with a known basis of popularity. In R, packages dimRed, dyndimred, intrinsicDimension.
- **Implementation:** mixture of R and C++ that are integrated by Rcpp. For numerical operations, RcppArmadillo is heavily used to take advantage of Armadillo C++ linear algebra library.
- **3 families of functions:** `do.{algorithm}`, `est.{algorithm}`, and `aux.{algorithm}` for DR, IDE, and auxiliary functions.
- Downloaded 1013 times per month on average from CRAN.
- **Future plan:** support for out-of-memory execution in response to the increased needs for big data analysis.

1.3 Example R Code

```
# Documentation example:
# do.idmap (Interactive Document Map)

library(Rdimtools)

## load iris data
data(iris)
set.seed(100)
subid = sample(1:150,50)
X = as.matrix(iris[subid,1:4])
lab = as.factor(iris[subid,5])
## let's compare with other methods
out1 <- do.pca(X, ndim=2)
```

```

out2 <- do.lda(X, ndim=2, label=lab)
out3 <- do.idmap(X, ndim=2, engine="NMP")
## visualize
opar <- par(no.readonly=TRUE)
par(mfrow=c(1,3))
plot(out1$Y, pch=19, col=lab, main="PCA")
plot(out2$Y, pch=19, col=lab, main="LDA")
plot(out3$Y, pch=19, col=lab, main="IDMAP")
par(opar)

```

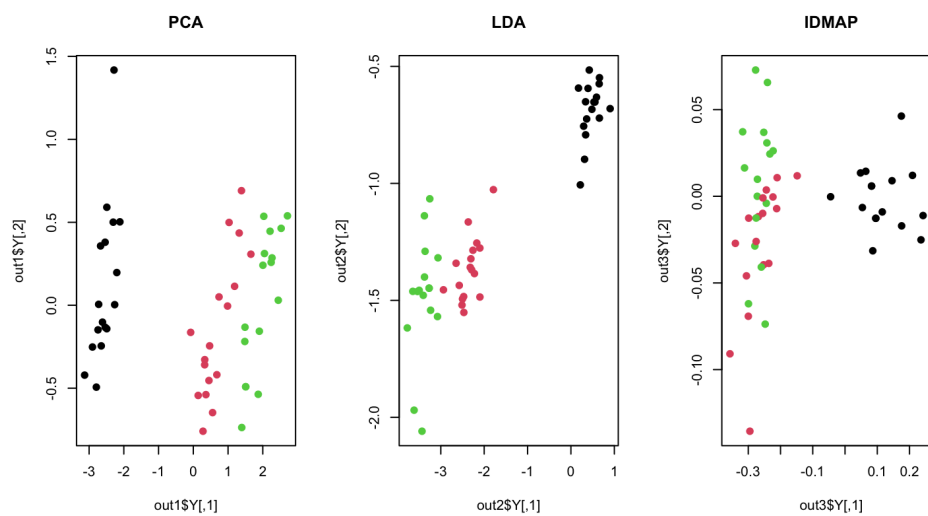


Figure 1: Rdimtools test output.

References

- [1] Kisung You and Dennis Shung. “Rdimtools: An R Package for Dimension Reduction and Intrinsic Dimension Estimation”. In: *Software Impacts* 14 (2022), p. 100414. ISSN: 26659638. DOI: 10.1016/j.simpa.2022.100414.