

# Package: **dobin** (via r-universe)

October 21, 2024

**Type** Package

**Title** Dimension Reduction for Outlier Detection

**Version** 1.0.4

**Maintainer** Sevvandi Kandanaarachchi <sevvandik@gmail.com>

**Description** A dimension reduction technique for outlier detection. DOBIN: a Distance based Outlier BasIs using Neighbours, constructs a set of basis vectors for outlier detection. This is not an outlier detection method; rather it is a pre-processing method for outlier detection. It brings outliers to the fore-front using fewer basis vectors (Kandanaarachchi, Hyndman 2020) <doi:10.1080/10618600.2020.1807353>.

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**Imports** dbscan, ggplot2, pracma

**RoxygenNote** 7.2.1

**Suggests** knitr, rmarkdown, OutliersO3, FNN

**VignetteBuilder** knitr

**Depends** R (>= 3.4.0)

**URL** <https://sevvandi.github.io/dobin/>

**Repository** <https://sevvandi.r-universe.dev>

**RemoteUrl** <https://github.com/sevvandi/dobin>

**RemoteRef** HEAD

**RemoteSha** d2724868b10e80bd670f2016a077802679059290

## Contents

autoplot.dobin . . . . .	2
dobin . . . . .	2

<b>Index</b>	<b>4</b>
--------------	----------

---

autoplot.dobin	<i>Plots the first two components of the dobin space.</i>
----------------	---

---

**Description**

Scatterplot of the first two columns in the dobin space.

**Usage**

```
## S3 method for class 'dobin'  
autoplot(object, ...)
```

**Arguments**

object	The output of the function 'dobin'.
...	Other arguments currently ignored.

**Value**

A ggplot object.

**Examples**

```
X <- rbind(  
  data.frame(x = rnorm(500),  
            y = rnorm(500),  
            z = rnorm(500)),  
  data.frame(x = rnorm(5, mean = 10, sd = 0.2),  
            y = rnorm(5, mean = 10, sd = 0.2),  
            z = rnorm(5, mean = 10, sd = 0.2))  
)  
dob <- dobin(X)  
autoplot(dob)
```

---

dobin	<i>Computes a set of basis vectors for outlier detection.</i>
-------	---

---

**Description**

This function computes a set of basis vectors suitable for outlier detection.

**Usage**

```
dobin(xx, frac = 0.95, norm = 1, k = NULL)
```

**Arguments**

<code>xx</code>	The input data in a dataframe, matrix or tibble format.
<code>frac</code>	The cut-off quantile for Y space. Default is 0.95.
<code>norm</code>	The normalization technique. Default is Min-Max, which normalizes each column to values between 0 and 1. <code>norm = 0</code> skips normalization. Other values of <code>norm</code> defaults to Median-IQR normalization.
<code>k</code>	Parameter <code>k</code> for <code>k</code> nearest neighbours with a default value of 5% of the number of observations with a cap of 20.

**Value**

A list with the following components:

<code>rotation</code>	The basis vectors suitable for outlier detection.
<code>coords</code>	The <code>dobin</code> coordinates of the data <code>xx</code> .
<code>Yspace</code>	The associated Y space.
<code>Ypairs</code>	The pairs in <code>xx</code> used to construct the Y space.
<code>zerosdcols</code>	Columns in <code>xx</code> with zero standard deviation. This is computed only if the number of columns are greater than the number of rows.

**Examples**

```
# A bimodal distribution in six dimensions, with 5 outliers in the middle.
set.seed(1)
x2 <- rnorm(405)
x3 <- rnorm(405)
x4 <- rnorm(405)
x5 <- rnorm(405)
x6 <- rnorm(405)
x1_1 <- rnorm(mean = 5, 400)
mu2 <- 0
x1_2 <- rnorm(5, mean=mu2, sd=0.2)
x1 <- c(x1_1, x1_2)
X1 <- cbind(x1,x2,x3,x4,x5,x6)
X2 <- cbind(-1*x1_1,x2[1:400],x3[1:400],x4[1:400],x5[1:400],x6[1:400])
X <- rbind(X1, X2)
labs <- c(rep(0,400), rep(1,5), rep(0,400))
dob <- dobin(X)
autoplot(dob)
```

# Index

`autoplot.dobin`, [2](#)

`dobin`, [2](#)