andrews.R

January 28, 2022

R topics documented:

---

andrews	*Compute Andrews curves*

Description

Compute data for Andrews plot, for visualizing clusters of multivariate data with ggplot2.

Usage

```
andrews(df, class_column, samples = 200, reorder = FALSE)
```

Arguments

df	Data frame to be used, should encompass column containing class names as well at least 2 numeric columns, preferably normalized to (0.0, 1.0).
class_column	String name of the column containing class names used for clustering.
samples	Number of points to plot in each curve.
reorder	Logical indicating whether to reorder numeric columns based on contributions to first principal component.

Details

Andrews curves have the functional form: $f(t) = x\_1/sqrt(2) + x\_2 \sin(t) + x\_3 \cos(t) + x\_4 \sin(2t) + x\_5 \cos(2t) + ...$ Where x coefficients correspond to the values of each dimension and t is linearly spaced between -pi and +pi. Each row of frame then corresponds to a single curve.

Examples

```
library(ggplot2)
df <- andrews(iris, "Species")
ggplot(df, aes(x = t, y = values, color = class_column, group = sample)) +
  geom_line() +
  scale_x_continuous(n.breaks = 7)
```

Index