The Journal of Open Source Software

Gala: A Python package for galactic dynamics

Adrian M. Price-Whelan[1, 2] **and Author 2**[2]

1 Lyman Spitzer, Jr. Fellow, Princeton University **2** Institution 2

DOI:

Software

- Review
- Repository
- Archive

Submitted:
Published:

Summary

Circular or directional statistics is the study of directional or angular data. Examples include data that is periodic in nature such as time (i.e. hours of the day, days of the year) and data that has no true zero such as compass bearing or wind direction, and circular statistics have been broadly applied across scientific disciplines. A common application in biological science is movement ecology, or the movement of organisms. Our software package extends principles used in movement ecology to personal location monitoring in order to decipher patterns in human activity.

Statement of need

`circleclust` is a collection of functions that facilitate the analysis of personal location data using R. Our algorithm detects changes in activity based on the angular variability, or circular variance, within a series of latitude and longitude coordinates. Like the variance statistic in linear data, circular variance is a measure of dispersion but is bounded between 0 and 1. A value of 0 indicates the sample directions (bearings) are uniform, while a value of 1 indicates a high degree of variability. The `circleclust()` function calculates circular varaince within a moving window and classifies coordinates as either 'static' or 'mobile' based on departures from a threshold value. Further, static points are categorized into distinct spatiotemporal clusters. The `circleclust()` function has been optomized to detect changes in activity pattern within a 5-minute moving window and a threshold circular variance of 0.7. However, these parameters can be modified.

Figure 1. P

`circleclust()` was developed to address a research gap in personal exposure studies and environmental epidemiology. Technological advances in mobile technology and consumer-grade sensors have expanded the ability of researchers to conduct personal monitoring studies at a large scale.

. Static coordinates are assigned a numeric cluster value that is unique by both space and time.

The package includes functions to aggregate data by time unit, impute missing location data, transform and visualize spatial data, and ascertain patterns in mobility.

To our knowledge, this if the first software package to be distributed publicly.

.`circleclust::circleclust()` algorithm

`circleclust` also includes

The underlying algorithm detects changes in activity based on the angular variability within a series of latitude and longitude coordinates. Less