Astronomical Data Analysis Software and Systems VIII
ASP Conference Series, Vol. 172, 1999
D. M. Mehringer, R. L. Plante, and D. A. Roberts, eds.

Large Scale Clustering in the Universe – A Package of Codes

Ivan Valtchanov

Institute of Astronomy, Bulgarian Academy of Sciences, 72 Tsarigradsko Chaussee Blvd, BG-1784 Sofia, Bulgaria

Abstract. A collection of codes is compiled for studying some aspects of the large scale clustering of galaxies, clusters and superclusters of galaxies.

1. Introduction

Measuring the degree of clustering is one of the most important tasks in the contemporary cosmology. Numerous methods for analysis and many theories were developed in the past forty years due to the efforts of many scientists (e.g., Peebles 1980, 1993).

Because the analysis of the large scale structure cannot be made into routine work, most of the steps for analysis require specific methods and construction of specific programs as well. That is why a package of codes for large scale structure analysis cannot be written as a general tool.

Nevertheless, a collection of procedures for studying some aspects of the clustering in the Universe can be constructed; inevitably any such package or collection of codes will only be a starting point and will require additional development. All the codes constituting the package are applicable to a large range of problems concerning the clustering of galaxies, clusters and superclusters.

2. Correlation functions

The correlation functions are very powerful instrument to study the large scale clustering of galaxies and clusters (e.g., Peebles 1980; Sicotte 1995). All the procedures implemented in the package require random catalog generations and these mock or Monte Carlo catalogs must contains all the selection effects presented in the real catalog.

The following selection effects are analyzed – distance, galactic latitude, galactic longitude, declination, supergalactic latitude.

2.1. Incorporating the selections

Each selection can be represented by a functional form and then it can easily be incorporated in the coordinate generation. Other possible variants where no assumption for the functional form for the selection were also implemented:

- Smoothing the cumulative distribution function (CDF) of the coordinate in consideration and using it as a template for random coordinate generation.
- Bootstrap resampling (Efron 1979).

We have used extensively all the methods for random catalogs creation and even combinations of them is possible. For example, we take uniform distribution of the galactic longitude, for the galactic latitude we implement the functional cosecant law and for the distance we use the bootstrap resampling.

2.2. Correlation function estimators and uncertainties

Depending on the dimensions of interest, we have developed the one dimensional (1D), 2D and 3D two-point correlation function, the cross-correlation function (Peebles 1980) and the correlation function for the line-of-sight and transverse direction pair separations $\xi(r_p, \pi)$ (e.g., Guzzo et al. 1997). An extensive use of the procedures can be found in Kalinkov, Valtchinov, & Kuneva (1998a, 1998b).

The following estimators were implemented in all the codes – RR estimator, DP estimator (Davis & Peebles 1983), H estimator (Hamilton 1993) and LS estimator (Landy & Szalay 1993).

We regard the following uncertainties for each estimator – Poissonian error, bootstrap error (Ling, Frenk, & Barrows 1986), bias-corrected bootstrap uncertainty (Efron & Tibshirani 1986).

3. Substructures searching

Many different methods are developed to search for substructures and comparisons to computer simulations are used for assessment of their performance (see for example Pinkney, Roettiger, & Burns 1996). In the package we have implemented a qualitative approach through the use of different mapping techniques – multiscale wavelet analysis (e.g., Slezak, Bijaoui, & Mars 1990; Starck, Murtagh, & Bijaoui 1998) and Gaussian adaptive and non-adaptive smoothing and filtering (e.g., Silverman 1986; Pisani 1993, 1996).

For the sake of quantifying the observed peaks in the density distribution, a method based on the work of Escalera & Mazure (1992) is implemented.

An application and results with the use of our procedures can be found in Kalinkov et al. (1998a).

4. Finding groups of objects

We have developed procedures based on the variants of the friends-of-friends (FOF) algorithm. Such procedures are used mainly for construction of catalogs of groups or superclusters of galaxies.

Huchra & Geller (1982), Nolthenius & White (1987) and Nolthenius (1993) procedures are implemented. A comprehensive comparison for the methods has been made by Frederic (1995).

A different approach for finding groups is developed by Kalinkov & Kuneva (1993). It is based on searching for groups of objects with chosen density contrast over the local density in given local volume. In a sense, this method is adaptive with respect to the local density and is also FOF method. The result from the application of this procedure is the construction of the largest catalog of superclusters of galaxies (Kalinkov & Kuneva 1993).

5. Void probability function

The void probability function (VPF) depends on the correlation function of all orders and in that aspect it is complementary to the correlation function; in some cases it also provides informative statistics (White 1979). An example of using the procedures described here can be found in Kalinkov et al. (1998a).

VPF in 2D/3D is the probability that a randomly thrown disk/sphere, with given area/volume in the distribution with mean surface/volume density, is empty of objects. The procedures implement the VPF definition with the including edge and selection effects.

6. Utilities

In this section various utility procedures are described. They are independent and can be used as a standalone procedures.

- An Interactive tool for robust estimation of location and scale (Beers, Flynn, & Gebhardt 1990)
- Cone diagrams
- Cosmological distances

7. Future Prospects

The future development of the package will be toward searching for different non spherical structures – filaments, sheets, walls. Searching for voids among galaxies, clusters and superclusters. Procedures for 3-point correlation function and correlation functions of higher order as well as power spectrum analysis and count-in-cell techniques are on the way. Other procedures for searching of groups, clusters and superclusters will be implemented in the package too. Refining some of the codes and making them more efficient for larger samples is an everyday work.

8. Implementation

The procedures requiring much CPU time are written in C (correlation functions, friends-of-friends algorithms) and where the analysis requires much visual inspection the procedures are written in IDL[1]. A few external procedures and functions are used in the C programs and they are taken from Numerical Recipes (Press et al. 1986) but replacing them with any other publicly available procedures is not a problem.

The IDL procedures use ASTROLIB library which is part of the library with many astronomical applications for IDL and is publicly available from `http://idlastro.gsfc.nasa.gov`. All the package with extensive documentation will be available from `http://www.astro.bas.bg/~ivan/lssp/` or by request.

[1]IDL is the trademark of Research Systems, Inc.

Acknowledgments. I am thankful to A. Pisani for providing me with his DEDICA package, T. Beers for ROSTAT package. I would like to thank H. Sicotte for providing his invaluable Ph.D. thesis work.

This work was supported by National Research Fund of the Bulgarian Ministry of Science, Education and Technology – contract F721/1997.

References

Beers T., Flynn K., & Gebhardt K. 1990, AJ, 100, 32

Davis M. & Peebles P. J. E. 1983, ApJ, 267, 465

Efron B. 1979, Ann. Stat., 7, 1

Efron B. & Tibshirani R. 1986, Stat. Sci., 1, 54

Escalera E. & Mazure A. 1992, ApJ, 338, 23

Frederic J. 1995, ApJS, 97, 259

Guzzo, L., Strauss, M. A., Fisher, K. B., Giovanelli, R., & Haynes, M. P. 1997, ApJ, 489, 37

Hamilton A. J. S. 1993, ApJ, 417, 19

Huchra J. & Geller M. 1982, ApJ, 257, 423

Kalinkov M. & Kuneva I. 1993, A&AS, 113, 451

Kalinkov M., Valtchanov I., & Kuneva I. 1998a, A&A, 331, 838

———, 1998b, ApJ, 506, 509

Landy S. D. & Szalay A. 1993, ApJ, 412, 64

Ling E. N., Frenk C. S., & Barrows J. D. 1986, MNRAS, 223, 21

Nolthenius R. 1993, ApJS, 85, 1

Nolthenius R. & White S. 1987, MNRAS, 235, 505

Peebles P. J. E. 1980, The Large Scale Structure of the Universe (Princeton: Princeton Univ. Press)

———, 1993, Principles of Physical Cosmology, (Princeton: Princeton Univ. Press)

Pinkney J., Roettiger K., & Burns J. O. 1996, ApJS, 104, 1

Pisani A. 1993, MNRAS, 265, 706

———, 1996, MNRAS, 278, 697

Press W., Teukolsky S., Vetterling W., & Flannery B. 1992, Numerical Recipes in Fortran (2nd ed.), (Cambridge: Cambridge Univ. Press)

Sicotte, H. 1995, Ph.D. thesis, Princeton Univ.

Silverman, B. 1986, Density Estimation for Data Analysis and Statistics (London: Chapman & Hall)

Slezak E., Bijaoui A., & Mars G. 1990, A&A, 227, 301

Starck, J.-L., Murtagh, F., & Bijaoui, A. 1998, Image Processing and Data Analysis (Cambridge: Cambridge Univ. Press)

White S. D. M. 1979, MNRAS, 186, 145