Cross Matching Galaxy Catalogs

Jaylen Luc, Cora Schallock

1 Abstract

Databases of different telescopes use different forms of galaxy indexing. SDSS uses 18 digit hashes to identify their galaxies. PanStarrs uses a variety of different catalogs that include Uppsala General Catalog (UGC), New General Catalog (NGC), Principal Galaxies Catalog (PGC), and Index Catalogs (IC) which are supplemental to the NGC. My task is to cross match PanStarrs galaxies within the Spin-Parity Catalog with galaxies with the sloan digital sky survey galaxies (SDSS) found in the Kelly-Final-GZ-all.tsv file. We want to find exactly what SDSS object ID matches with any given galaxy in the Spin-Parity only using Right Ascension (RA) and Declination (Dec).

2 Methodology

Stars do not change positions – their RA or Dec– when the Earth rotates or orbits around the Sun. The Altitude and Azimuth will change relative to where the Earth is in its orbit but the RA and Dec will not change in the celestial sphere. But using just RA and Dec is not a robust modus operandi of determining uniqueness of a celestial object. Stars change their position relative to the center of mass of the solar system called proper motion and the orientation of the Earth's North Pole due to precession. Positions of the stars can also change due to Mutations. There are small deviations in the RA and Dec between different telescopes and in different catalogs. If there are potential marginal RA and Dec differences, it is sufficient to assume that if we calculate the closest match objects and then manually check the matches or do image spectroscopy, we can say with relatively high confidence that they are the same galaxy. We check the angular distance, not the euclidean distance, between the two galaxies because if we know their RA and Dec, we know their angle. So if we take the cosine of the angular distance between the galaxies the equation would be :

$$\cos(C) = \sin(d_a)\sin(d_b) + \cos(d_a)\cos(d_b)\cos(r_a - r_b)$$

Figure 1: *angular cosine distance formula. d_a and d_b is the declination of galaxies a and b. R_a and r_b are the Right Ascensions of galaxies a and b.*

RA and Dec for this cosine function is in radians. The algorithm converts all degrees into radians, puts all the ras, decs into a 2-dimensional Numpy array, puts the corresponding identifier into another array with its index corresponding to the same galaxy in the RA Dec 2-dimensional array. It then computes for every galaxy in the spin-parity catalog, all the closest matched SDSS galaxies. It is performed by applying the cosine distance to all other SDSS galaxies and then taking the maximum cosine distance. In theory, all galaxies in the Spin-Parity should have one unique closest match. The program is O(N^2) because for each galaxy in Spin-Parity it must be calculated with every other SDSS galaxy. This is not ideal; however, we can accelerate python by exploiting Numpy arrays by utilizing the GPU. We essentially just treat the cosine function for every iteration as a matrix multiplication where d_a and r_a are constant for one iteration but d_b and r_b are a 1 x N matrix. A log is returned with all the closest matches and all relevant information.

3 Todo

We can employ other methods such as triangular tessellation to index galaxies and find close matches , use spectrography to verify closest matches, or accelerate hardware to speed up crossmatches for bigger data sets.