Astronomy Problem Statement

Date:25/03/2022

1 Problem Statement

Astrometric analysis of open clusters is a valuable tool for understanding the formation and evolution of our galaxy. Open clusters are groups of stars that formed from the same molecular cloud, and their properties can provide insight into the conditions and processes of star formation. By measuring the positions and motions of individual stars within an open cluster, astrometric analysis can help to determine its age, distance, and kinematics. This information can then be used to study the chemical and dynamical evolution of the cluster and the larger region of the galaxy in which it resides. Additionally, astrometric analysis of open clusters can be used to test models of stellar evolution and to study the effects of interactions between stars on the evolution of the cluster. Overall, astrometric analysis of open clusters is a powerful tool for studying the structure and evolution of our galaxy and the processes that give rise to the stars within it.

1.1 Objective

The objective of this exercise is to give you a flavour of how research in astronomy could look like. If you are interested you can study this article for getting a more detailed information.

You will have to do the following things:

1. Use Gaussian Mean Mixture Model, to seperate out the stars belonging to the open cluster from the background stars
2. Create a quiver plot to visualise the proper motion of stars in the Right Ascension and Declination (RA-Dec) plane
3. Plot a Colour Magnitude Diagram of the member stars

The data of the stars has been downloaded from Gaia Archive and was obtained through Gaia Space Telescope launched by Europeon Space Agency.

2 Gaussian Mean Mixture Model

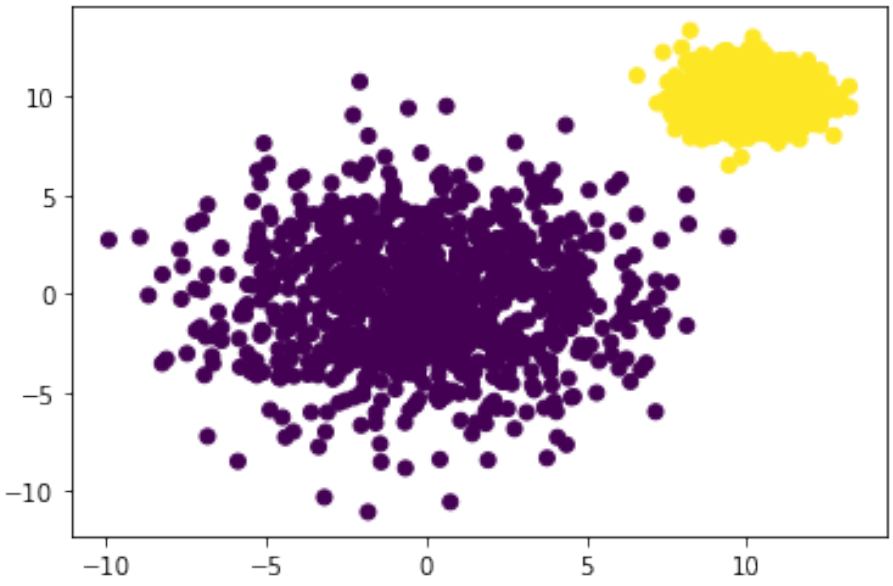
The Gaussian Mixture Model (GMM) is a statistical model that is often used for clustering and density estimation in data analysis. It assumes that the data is generated from a mixture of several Gaussian distributions, each with a different mean and variance.

Imagine you have a dataset and want to group similar data points together. The GMM helps you do this by modeling the data as a combination of several Gaussian distributions with different means and variances. Each data point is then assigned to one of these Gaussian distributions based on its probability of belonging to each distribution.
Now we will use it to seperate out stars belonging to the open cluster from the background stars

Listing 1: Example code of finding Gaussian Mean Mixture for generated data

```
from sklearn.mixture import GaussianMixture
import numpy as np

# generate some sample data
n_samples = 1000
cluster_mean = [10, 10]
noise_mean = [0, 0]
cluster_cov = [[1, 0], [0, 1]]
noise_cov = [[10, 0], [0, 10]]
X_cluster = np.random.multivariate_normal(cluster_mean, cluster_cov, n_samples)
X_noise = np.random.multivariate_normal(noise_mean, noise_cov, n_samples)
X = np.vstack((X_cluster, X_noise))

# fit the GMM
gmm = GaussianMixture(n_components=2, covariance_type='full')
gmm.fit(X)

# predict the cluster membership of each data point
y_pred = gmm.predict(X)

# plot the data points and their predicted cluster membership
import matplotlib.pyplot as plt
plt.scatter(X[:, 0], X[:, 1], c=y_pred)
plt.show()
```

Result of the above example code

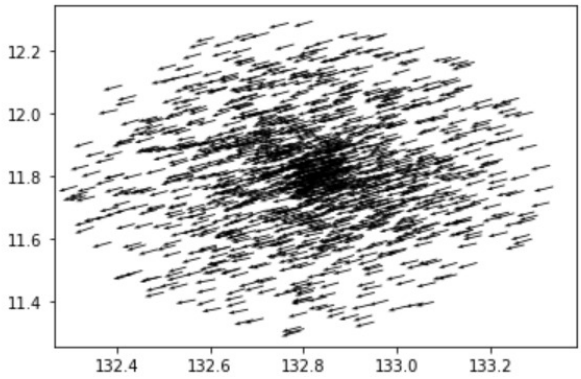


In the results we can see that using GMM we are able to determine which stars belong to which cluster. Use this method to extract the indices of stars belonging to the open cluster.

3 Proper Motion Analysis and Plotting Colour Magnitude Diagram

To loosely verify if all the stars that we have selected using GMM actually belong to the open cluster we can create a quiver plot that plots the position (RA will be on X-axis and DEC will be on Y-axis) of the star and its proper motion as its velocity (U: propermotion in right ascension (pmra) and V: propermotion in declination (pmdec))

The results should look as follows



To create the colour magnitude diagram you simply need to plot the "bp_rp" versus "phot_g_mean_mag" on the Y-axis. The results would be as follows: