Membership determination of open cluster with parametric method: cross entropy

Itsna Khoirul Fitriana and M. Ikbal Arifiyanto

1 Astronomy Program, Institut Teknologi Bandung, Indonesia

E-mail: itsnafitriana@gmail.com

Abstract. Stellar open clusters are useful for studying the structure and evolution of Galaxy. Determination of the open cluster members can be viewed from their kinematics: proper motion and radial velocity. It is because the star cluster members will have a motion in the same direction, towards a convergent point. Parametric method with a double elliptic bivariate gaussian function which contains 11 parameters is used as a model of the proper motion distribution. The value of these parameters are determined numerically by cross entropy method. This method is a method of data fitting using the likelihood function whose value is maximized to get the proper parameters. This method solves the problem of determining the initial value that is often difficult to determine in other numerical methods. There are four open clusters which we have analysed: NGC 2244, ASCC 100, NGC 5168, and NGC 2169. The resulting of proper motion parameters of this method matches the parameters in the literature.

1. Introduction

Study of structure and evolution of Galaxy can be complete by studying open clusters. Before performing further analysis of the open cluster, the membership determination of an open cluster is a process that should not be missed. This research aims to determine the membership of some open clusters with double elliptic bivariate gaussian function which contain of 11 parameters by [5]. The parameters in the equation are solved numerically using cross entropy (CE) method that contains of maximum likelihood function method. Actually there are many other methods to solve the equation, but the CE method is chosen because it has many advantages. The CE method uses the initial value in a range, it can answer difficult issues of other methods in determining the initial value.

2. Data

We used kinematic data (position, proper motion, and magnitude) of each cluster from different sources. We take data of NGC 2244 by [3]. This object is the benchmark model for this research and [5] which use same function. The other clusters, ASCC 100, NGC 5168, and NGC 2169 are taken from from UCAC4, the Fourth US Naval Observatory CCD Astrograph Catalog. We choose another object as the sample of an open cluster which has stars of field less than the stars belong to cluster (ASCC 100) and an open cluster which has field stars more than the stars belong to cluster (NGC 5168). The fourth object, NGC 2169, is the object whose data taken with twice the radius which is written in Karchenko atlas page [2]. These objects had been determined by [1] using another gaussian model which solved numerically by CE method. The fourth cluster is used to check whether tour
method can be used for a variety of data. Then, it is used on other clusters that haven’t been checked before.

3. Method
The method used for the four clusters is similar. The method is CE method which has same steps with [4]. First, the determination of the initial value range using mean and dispersion value of proper motion data [5]. Then, proper motion data is fitted with double elliptic bivariate gaussian function (equation 1) using CE method. Then, the best value of the parameters are used to calculate the membership probability of each stars with Bayessian model. We determine the stars belong to cluster which has probability more than equal to 0.5. Then the cluster members are plotted in a HR diagram to see the distribution of evolutionary track (figure 1). The last step is to compare the result with other literature. We compare the probability of the stars and the value of cluster’s parameters in D value (equation 2). When the D value is close to 0, it means the comparison is fit.

4. Result and discussion

4.1 Results
The results of this research are:
- The model and method can be applied for the four open clusters: NGC 2244, ASCC 100, NGC 5168, and NGC 2169. The best parameter values are listed in the table 1

| Parameter | NGC 2244   | ASCC 100    | NGC 5168    | NGC 2169    |
|-----------|------------|-------------|-------------|-------------|
| $N_f$     | 0.50 ± 0.28| 0.30 ± 0.16 | 0.52 ± 0.29 | 0.36 ± 0.22 |
| $\sigma_xc$ | 0.09 ± 0.07| 3.06 ± 1.77 | 2.65 ± 1.56 | 1.03 ± 0.58 |
| $\sigma_{yc}$ | 0.09 ± 0.06| 2.88 ± 1.62 | 2.53 ± 1.46 | 0.95 ± 0.54 |
| $\rho_f$  | 0.02 ± 0.11| 0.01 ± 0.14 | -0.01 ± 0.16| 0.00 ± 0.13 |
| $\mu_{xf}$| -0.04 ± 0.03| -2.58 ± 1.58| -3.44 ± 1.96| -2.63 ± 1.47|
| $\mu_{yf}$| -0.06 ± 0.05| -2.14 ± 1.38| -3.41 ± 1.97| -2.75 ± 1.60|
| $\sigma_{xf}$| 0.77 ± 0.46| 20.83 ± 11.75| 30.77 ± 17.91| 22.68 ± 12.73|
| $\sigma_{yf}$| 0.78 ± 0.45| 20.59 ± 11.54| 30.30 ± 17.13| 22.93 ± 12.77|
| $\rho_c$  | 0.01 ± 0.12| 0.04 ± 0.14 | -0.03 ± 0.19 | 0.01 ± 0.16 |
| $\mu_{xc}$| 0.01 ± 0.01| 0.02 ± 0.07 | -3.96 ± 2.24 | -7.21 ± 4.74 |
| $\mu_{yc}$| 0.02 ± 0.01| -2.41 ± 1.34| -4.05 ± 2.30 | -1.48 ± 0.84 |

- D value between the research and first literature [5] for NGC 2244 and [1] for others) is more close to zero than D value of second comparison (literature by [2]). It can be seen in the table 2.
Table 2. D values for each cluster, D1 shows a comparison of the value of D to the data [1] except for NGC 2244 with the data [5], while D2 shows the difference with [2]

|       | NGC 2244 | ASCC 100 | NGC 5168 | NGC 2169 |
|-------|----------|----------|----------|----------|
| Nilai D1 | 1.13     | 0.43     | 0.02     | 0.24     |
| Nilai D2 | 1.40     | 2.52     | 2.50     | 0.21     |

4.2 Discussion
The results show the successful performance of the model and method. So, it can be used to other open clusters which haven’t been checked before. It will help understand our Galaxy. But in practice there are still some steps that can be done in the future research. It aims to make research results getting better. Some suggestions put forward are as follows: (i) Radial velocity factor can be added in the membership determination. The study is conducted in three dimensions by adding a radial velocity component in addition to motion. (ii) Comparison between parametric and non-parametric method will be good to do to make sure the better method.

Appendices

\[ \Phi(\mu_{x1}, \mu_{y1}) = N_f \Phi(\mu_{x1}, \mu_{y1}) + (1 - N_f) \Phi(\mu_{x1}, \mu_{y1}) \]
\[ = \frac{N_f}{2\pi \sigma_f \sqrt{1 - \rho_f^2}} \exp \left( -\frac{1}{2(1-\rho_f^2)} \left( \frac{\mu_{x1} - \mu_{x\text{f}}}{\sigma_{x\text{f}}} \right)^2 + \left( \frac{\mu_{y1} - \mu_{y\text{f}}}{\sigma_{y\text{f}}} \right)^2 \right) \]
\[ + \frac{(1 - N_f)}{2\pi \sigma_c \sqrt{1 - \rho_c^2}} \exp \left( -\frac{1}{2(1-\rho_c^2)} \left( \frac{\mu_{x1} - \mu_{x\text{c}}}{\sigma_{x\text{c}}} \right)^2 + \left( \frac{\mu_{y1} - \mu_{y\text{c}}}{\sigma_{y\text{c}}} \right)^2 \right) \]

where \( N_f \) is a proportion of field to the number of stars in total, \( \mu_{x1} \) and \( \mu_{y1} \) are proper motion data in x and y directions, \( \sigma_{x\text{f}} \) and \( \sigma_{y\text{f}} \) are the standard deviation in the x and y directions for the field stars, \( \mu_{x\text{f}} \) and \( \mu_{y\text{f}} \) are proper motion parameter in x and y directions for the field stars, \( \mu_{x\text{c}} \) and \( \mu_{y\text{c}} \) are proper motion parameter in x and y directions for cluster stars, \( \sigma_{x\text{c}} \) and \( \sigma_{y\text{c}} \) are the standard deviation in the x and y directions for the cluster stars, \( \rho_f \) and \( \rho_c \) are the coefficient of rotation for the field and cluster stars.

\[ D = \frac{(\mu_{x\text{cE}} - \mu_{x\text{LIT}})^2}{(\sigma_{x\text{cE}})^2 + (\sigma_{x\text{LIT}})^2} + \frac{(\mu_{y\text{cE}} - \mu_{y\text{LIT}})^2}{(\sigma_{y\text{cE}})^2 + (\sigma_{y\text{LIT}})^2} \]

Where CE declared value results of Cross Entropy method which we did and LIT stated value parameter of literature which we compare.
Figure 1. Color Magnitude Diagram (CMD) refer to diagram of Hertzsprung-Russell (HR) for the stars whose probability value $\geq 0.5$. Color dots indicate the class of probability with 0.1 width of probability value: $0.5 \leq p_6 \leq 0.6$; $0.6 < p_7 \leq 0.7$; $0.7 < p_8 \leq 0.8$; $0.8 < p_9 \leq 0.9$; $0.9 < p_{10} \leq 1.0$

References

[1] Dias W S et al. 2014 A&A 564 A79
[2] Kharchenko N V, Piskunov A E, Röser S, Schilbach E, Scholz R D 2015 A&A 438 1163K
[3] Marschall L, Van Altena W, and Chiu L 1982 AJ 87 1497
[4] Monteiro H, Dias W S, Caetano T C 2010 A&A 516 A2
[5] Sabogal-Martnez B A, Garca-Varela J A G, Higuera M A, Uribe A, Brieva E 2001 Revista Mexicana de Astronoma y Astrofisica 37 105