Activity 15 Expectation Maximization

Kyngzer-Rem Vargas*

National Institute of Physics, University of the Philippines Diliman, Quezon City, 1101
*Corresponding author: kmvargas2@up.edu.ph

Objective(s)

1. Implement the expectation maximization (EM) algorithm to the fruits data.
2. Overlay a plot of the estimated pdf (d-dimensional Gaussian distribution) on the fruit feature space [1, 2].

Results

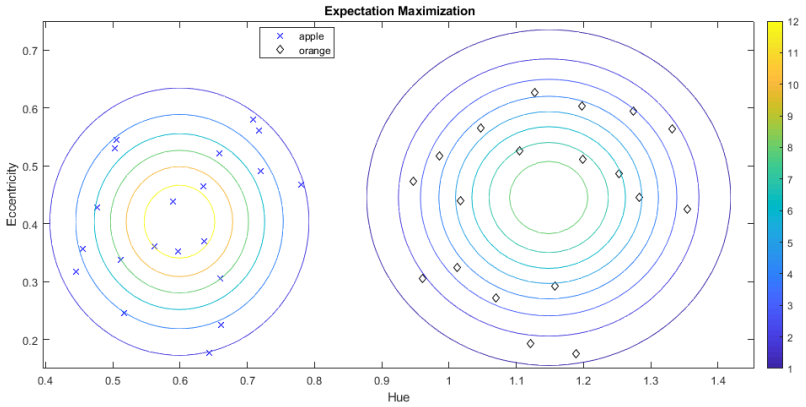

Figure 1: Estimated pdf on the fruit feature space.

Comment(s)

Expectation Maximization. Figure 1 shows that the EM algorithm was able to group the two different classes into clusters. The multiple circles and/or ellipses shows the effect of the E-step to calculate the log likelihood of the data cluster/class. The contour was plotted with the use of the co-variance of each feature for each class which yields the d-dimensional Gaussian distribution.

Self-Evaluation

I would rate myself a 10. The objectives for this activity was met. The data was successfully clustered by being able to group the two data classes. The updating rules were properly implemented onto the fruit feature space and the stopping condition that was used was until reached a certain number of iterations was reached. Besides the activity manual provided, I also read up on multiple websites as reference for the algorithm to get a better understanding and in particular these helped me the most [3, 4].

References

[1] M. Soriano, A15 - expectation maximization 2019.pdf.
[2] M. Soriano, Machine learning intro.pdf.
[3] TheMathWorksInc., fitcnb (2006), last accessed 19 November 2019, https://uk.mathworks.com/help/stats/fitcnb.html.

[4] B. Klein, Unsupervised learning: Clustering: Gaussian mixture models (gmm) (2017), last accessed 19 November 2019, https://www.python-course.eu/expectation_maximization_and_gaussian_mixture_models.php.