

Probabilistic Clustering

Probabilistic clustering (aka Gaussian Mixture Model)

- Unsupervised machine learning technique
- Divides data into cluster
- What is the right number of cluster?

Gaussian Mixture Model (GMM) algorithm

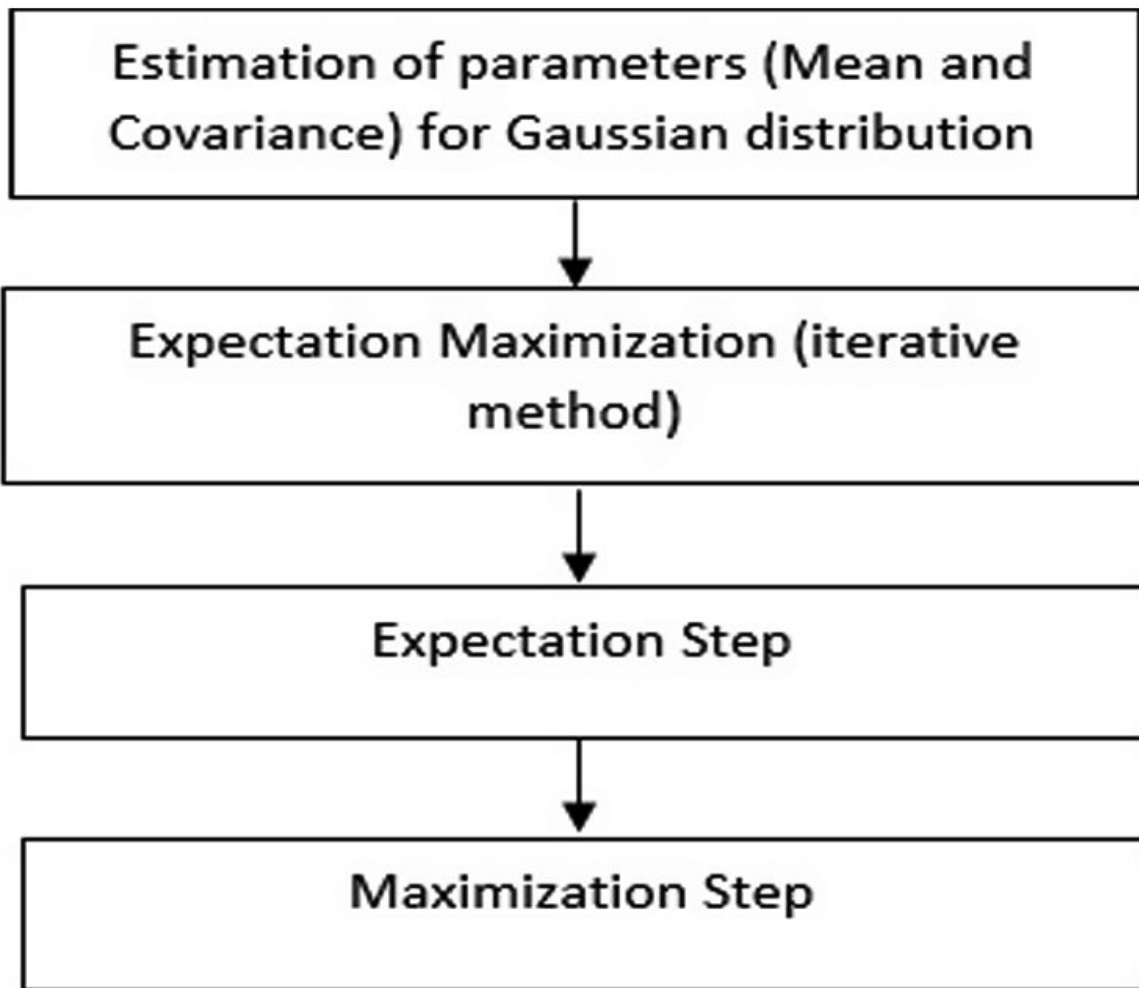
- Can get probability associated with each cluster for given cluster
- Algorithm consist of two primary steps :
 - a. Expectation
 - b. Maximization

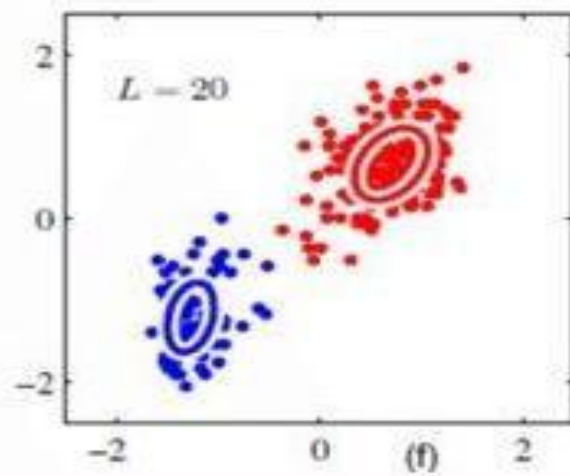
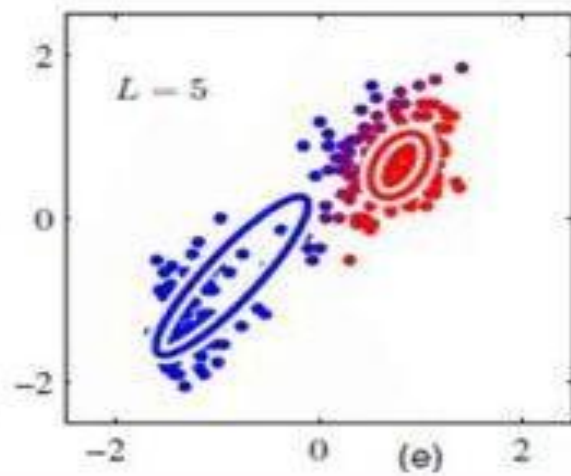
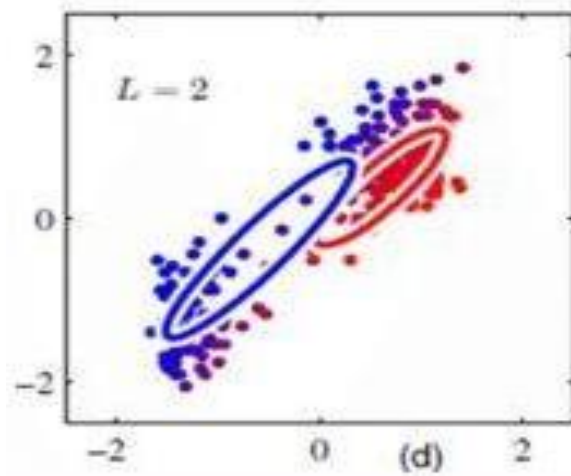
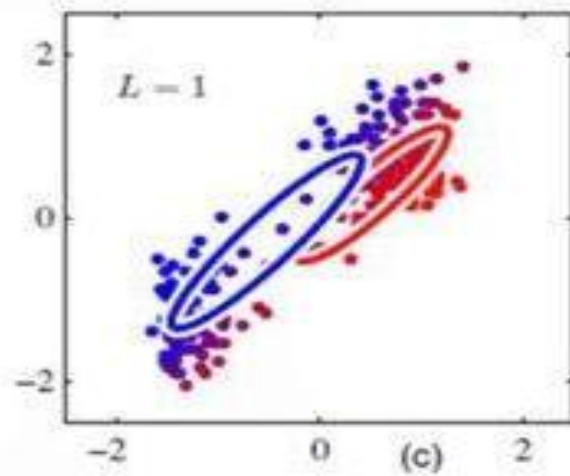
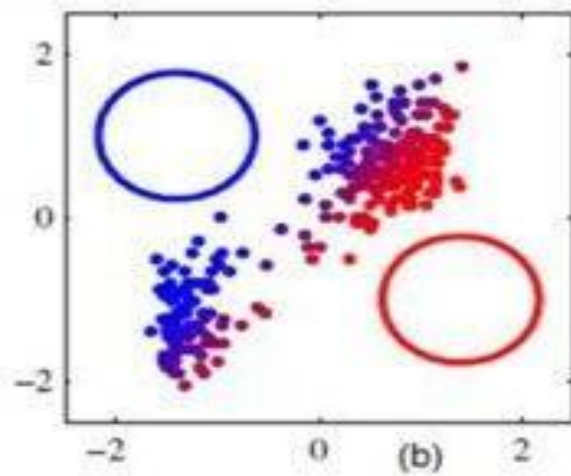
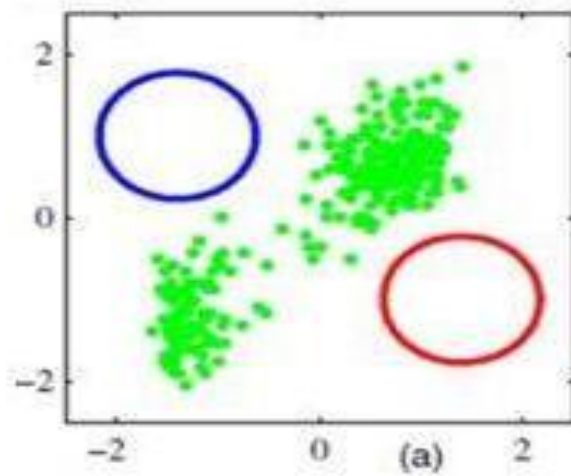
GMM Algorithm (Cont...)

1. **Expectation** - for each data point 'p' calculate that what is the probability that point 'p' belongs to cluster 'c'.
2. **Maximization** - adjust the mean, variance, and weight of each data point to maximize the possibility

GMM Flow chart

EM algorithm for
GMM





Reference

[1] Image - GMM Algorithm

https://media.springernature.com/original/springer-static/image/chp%3A10.1007%2F978-981-13-7403-6_1/MediaObjects/469174_1_En_1_Fig3_HTML.png

[2] Image - GMM $L=1$ to $L=20$

<https://www.analytixlabs.co.in/blog/wp-content/uploads/2020/07/image-14-1-600x450.jpg>