Semester Project: Gaussian Mixture Model

Xiyou Zhou, 13307130189
Computer Science and Technology

July 1, 2017

1 Gaussian Mixture Model with EM Algorithm

I start with plotting with 3 different Gaussian Distributions.

One is at $\mu = (0,0), \Sigma = \begin{bmatrix} 50 & 20 \\ 30 & 40 \end{bmatrix}$ with 2500 points, one is at $\mu = (50, 40), \Sigma = \begin{bmatrix} 400 & 500 \\ 200 & 400 \end{bmatrix}$ with 3000 points, and the last one is at $\mu = (-10, 40), \Sigma = \begin{bmatrix} 100 & 20 \\ 60 & 40 \end{bmatrix}$ with 4500 points.

The initial parameters are set with a KMeans result of 3 kernels. After 30 epochs, the distribution of three clusters changed from 30%, 25% and 45% to 25%, 30% and 45%. The kernel location changed from (4.30, 0.99), (54.54, 46.57), (-9.77, 40.01) to (0.10, -0.07), (49.79, 39.93), (-9.77, 40.05).

2 Gaussian Mixture Model with Gibbs Sampling

Not finished. Will implement later in the summer.

3 Analysis

Complex models generally leads to more accurate results and more data can help generate more accurate estimation. EM Algorithm and Gibbs Sampling both can be applied to latent parameter estimation. However, EM Algorithm always maximizes the result of certain likelihood function while Gibbs Sampling can be constructed from Monte-Carlo Markov Chain and its convergence is also guaranteed by MCMC.

EM as a maximization/maximization method, Gibbs as a variation of Generalized EM. Instead of maximizing at each of these two steps(E, M), Gibbs algorithm use the conditional distributions to sample from them.

References

[1] New York University: Mixture Models EM algorithm, Lecture 21
    http://cs.nyu.edu/~dsontag/courses/ml12/slides/lecture21.pdf

[2] C19: Lecture 4: A Gibbs Sampler for Gaussian Mixture Models
    http://www.robots.ox.ac.uk/~fwood/teaching/C19_hilary_2013_2014/gmm.pdf

[3] Gibbs and Metropolis sampling (MCMC methods)
    http://www.cs.cmu.edu/~tom/10-702/GibbsAndMCMCsampling.pdf
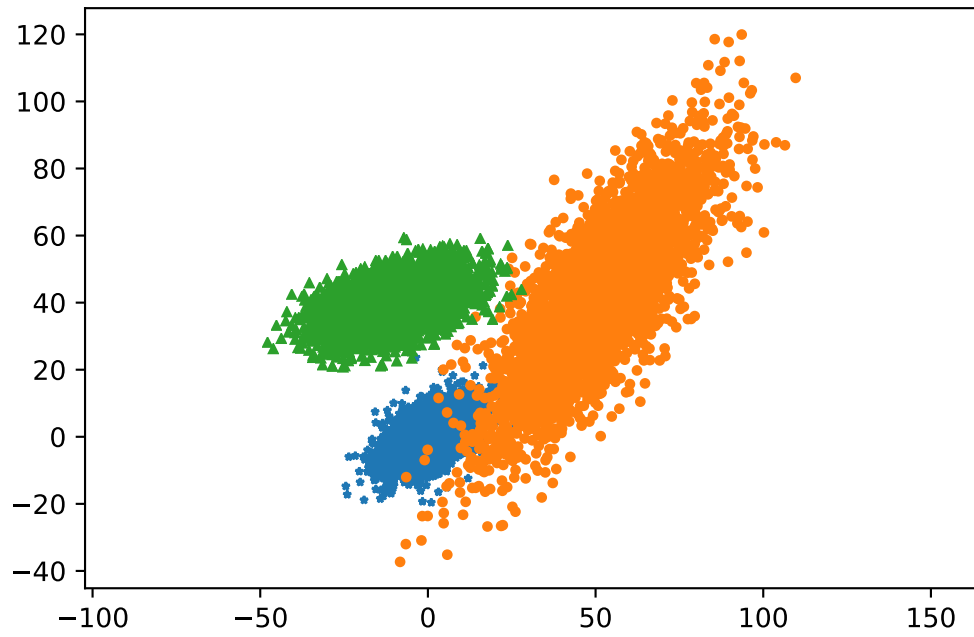
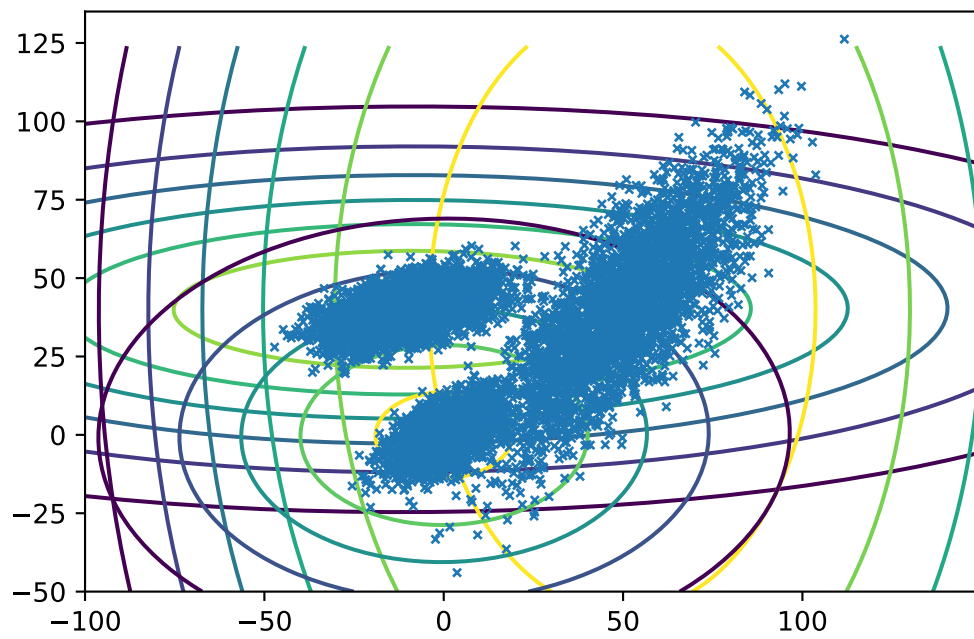## 4 Appendix

Figure 1: Plot of original distribution of all points

Figure 2: Plot of EM Algorithm fit result after 30 epochs