Homework 4 Group Report

**(Xingyu Huang, Chunlei Wang, Wenbo Zhao, and Soyeon Park)**

**Result:**

**K = 1:**

Step 2:

Mu = [15.481698203115867]

Sigma(variance) = [1.0]

Alpha = [67.51493830369019]

Log-likelihood = -182784.35282072006

Step 3:

Mu = [15.481698203115867]

Sigma(variance) = [1]

Alpha = [67.51493830369019]

Log-likelihood = -182784.35282072006

**K = 3:**

Step 2:

Mu = [5.5092793923195496, 25.486654429321, 15.44916078770705]

Sigma(variance) = [0.3333333333333333, 0.3333333333333333, 0.3333333333333333]

Alpha = [1.0302576351879733, 0.9980966180379126, 0.9671159482381272]

Log-likelihood = -11215.271295798197

Step 3:

Mu = [5.509279396102553, 25.486654429387666, 15.449160789903418]

Sigma = [1, 1, 1]

Alpha = [1.0302576590926606, 0.9980966179932715, 0.9671159435031166]

Log-likelihood = -8520.167721063615

**K = 5:**

Step 2:

Mu = [4.647662596490959, 25.486654429321, 16.230988210634095, 6.269489296892197, 14.69647741686372]

Sigma(variance) = [0.15643042151843675, 0.3333333333333333, 0.16363796197507682, 0.1766583053900404, 0.16973630110816193]

Alpha = [0.38814614308563333, 0.9980966180379126, 0.35451081806322526, 0.3651151217627762, 0.4017232839295901]

Log-likelihood = -11728.637068746486

Step 3:

Mu = [5.330976547551546, 25.48665442951566, 15.45031578099601, 5.6934006489394395, 15.448006029011324]

Sigma(variance) = [1, 1, 1, 1, 1]

Alpha = [1.0133551031675254, 0.9980966175354518, 0.9670227261399944, 0.9809818360810949, 0.967206469067667]

Log-likelihood = -5812.384648968671

What method did you use for initializations? Is the result sensitive to the initial values?

I initialize the mu, sigma and alpha by using python’s random function.

For mu: it will randomly generate K numbers between maximum of data and minimum of data.

For sigma: it will randomly generate K numbers between maximum of data and minimum of data.

For alpha, it will randomly generate K numbers between 0 and 1 and make sure summary of alpha equal to 1;

Of course, the result sensitive to the initial values, the difference shows between step 2 and step 3.

What is the true value of K? Why?

I think the true value of K is 3. Because for the Step 2. We can find out the log-likelihood from K=3 is closest to 0, and all variable from alpha is also closest to 1 too. So we think K=3 is the true value.