

Machine Learning Exercise 6

Florian Schneider, Amin Kasipour

03.12.12

1 Generated Data

In total $N = 300$ data points were generated using three Gaussians with the parameters shown in Table 1. The means were chosen by hand where as the covariance matrices were generated randomly. The generated data points are depicted in Figure 1.

2 Gaussian Mixture Model

In the documented run, the GMM converged after 35 epochs with the parameters described in Table 2: The resulting clustering and the likelihood over epochs are illustrated in Figure 2 and Figure 3 respectively:

3 k-Means Clustering

After 10 epochs k-Means with $k = 3$ resulted in the centroids shown in Table 3. The clustering is illustrated in Figure 4.

k	μ	Σ
0	$\begin{pmatrix} 2.0 \\ 2.0 \end{pmatrix}$	$\begin{pmatrix} 0.17469765 & 0.66551677 \\ 0.47888327 & 0.72910767 \end{pmatrix}$
1	$\begin{pmatrix} -2.0 \\ -2.0 \end{pmatrix}$	$\begin{pmatrix} 0.39959492 & 0.43382484 \\ 0.68072252 & 0.36495781 \end{pmatrix}$
2	$\begin{pmatrix} 0.0 \\ 0.0 \end{pmatrix}$	$\begin{pmatrix} 0.17371089 & 0.33020753 \\ 0.96190678 & 0.20638108 \end{pmatrix}$

Table 1: Parameters of the Normal Distributions to generate the data points

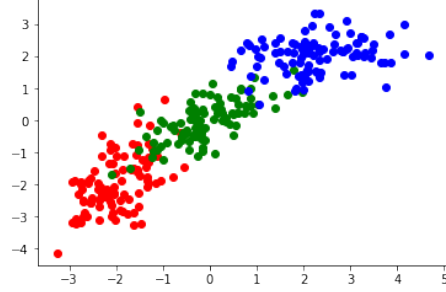


Figure 1: 300 data points generated by Normal Distributions with parameters of Table 1

\mathbf{k}	μ	Σ	ϕ
0	$\begin{pmatrix} 2.24932261 \\ 2.04490088 \end{pmatrix}$	$\begin{pmatrix} 0.79501829 & 0.09341968 \\ 0.09341968 & 0.27997369 \end{pmatrix}$	0.33385
1	$\begin{pmatrix} -1.94410665 \\ -1.86806968 \end{pmatrix}$	$\begin{pmatrix} 0.35521996 & 0.30531853 \\ 0.30531853 & 0.79764169 \end{pmatrix}$	0.34466
2	$\begin{pmatrix} -0.06559494 \\ -0.01846601 \end{pmatrix}$	$\begin{pmatrix} 0.50113652 & 0.30644989 \\ 0.30644989 & 0.39121782 \end{pmatrix}$	0.32147

Table 2: Final model parameters of the GMM with k=3 normal distributions

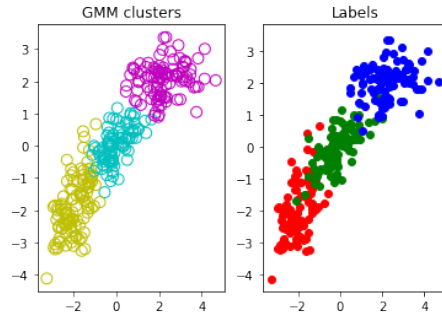


Figure 2: Final GMM clusters

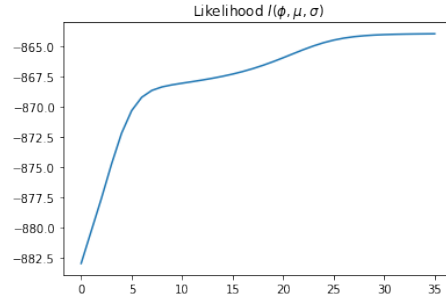


Figure 3: Likelihood over epochs

\mathbf{k}	μ
0	$\begin{pmatrix} 2.27271822 \\ 2.0261508 \end{pmatrix}$
1	$\begin{pmatrix} -1.93579551 \\ -1.92003393 \end{pmatrix}$
2	$\begin{pmatrix} -0.06199018 \\ 0.09040654 \end{pmatrix}$

Table 3: asd

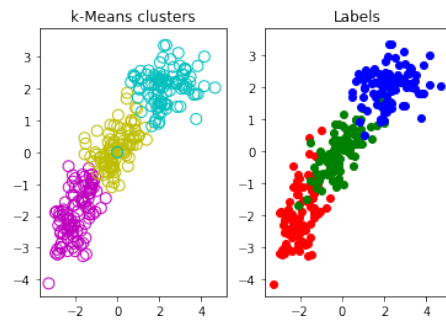


Figure 4: Final clustering of k-Means