COURSEWORK

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

Advanced Statistical Machine Learning

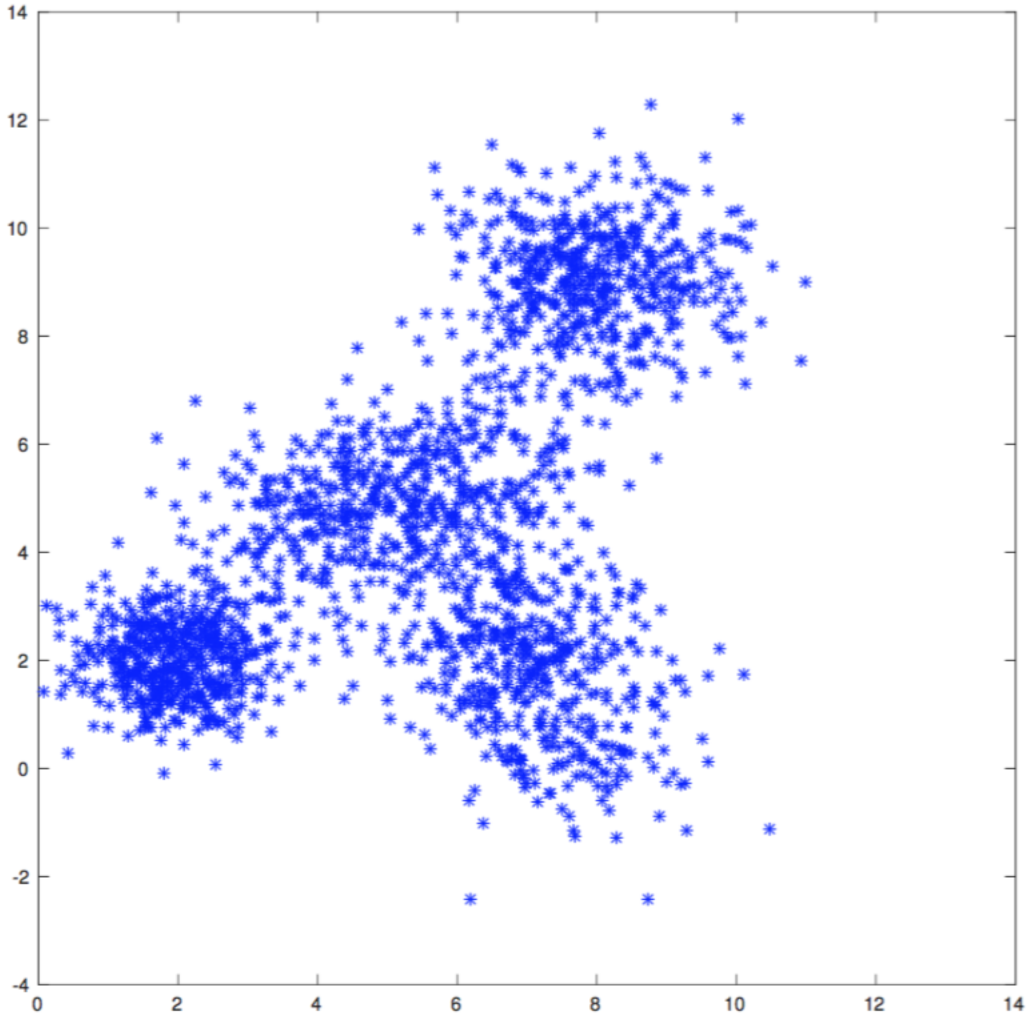
Coursework 1

Author:
Pavan Pinnaka

Date: February 10, 2017

1 Question 1

1.1 From left to right in Fig.1 and without doing any calculations, what is the form of the covariance matrix for each Gaussian (i.e., is it an isotropic1 or an anisotropic one)? Why? Briefly explain.



The covariance matrix defines the shape of the distribution. An isotropic covariance represents a spherical data cloud. Visually without any calculations, we can spot 4 different clusters within the data set. The first cloud on the bottom left should have an isotropic covariance matrix since the data points within that cluster form a perfect sphere. The second one (going right) seems to be representing a positive correlation between the two variables. Thus the covariance matrix isn't isotropic. The third one shows a negative correlation and thus isn't isotropic and finally the last one (top right cluster) again exhibits a spherical shape, thus has an isotropic covariance matrix.

2 Question2

2.1 Consider a special case of a GMM in which the covariance matrices Σ_k of the components are all constrained to have a common value Σ. Derive analytically the EM equations for maximising the likelihood function under such a model

$$\begin{aligned}G(\Theta) &= \sum_{n=1}^{N}\sum_{k=1}^{4} E_{p(Z|X,\Theta)}(z_{nk})\left[\ln(N(X_n|\mu_k,\Sigma)+\ln(\pi_k)\right] \\ &= \sum_{n=1}^{N}\sum_{k=1}^{4} \gamma(z_{nk})\left[-\frac{1}{2}(x_n-\mu_k)-\frac{1}{2}(F\ln(2\pi)+\ln(|\Sigma_k|))+\ln(\pi_k)\right]\end{aligned} \tag{1}$$

Now since we know that $\Sigma_k = \Sigma$, we can say (Note we set K=4 since we are considering 4 clusters in this example):

$$\begin{aligned}\frac{\partial G(\Theta)}{\partial \Sigma} &= 0 \\ \Sigma &= \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{4}\gamma(z_{nk})(x_n-\mu_k)(x_n-\mu_k)^T\end{aligned} \tag{2}$$

Resulting PDFs for the original data set after 10 iterations: