

EM算法类似于kmeans算法，但是EM算法其实包括了Kmeans算法

首先初始步骤分为两个，先随机选取k个点作为簇类中心，之后为每一个簇类设定参数，包括均值，和协方差矩阵，每一个簇类的先验概率。均值代表簇的中心，协方差矩阵代表簇类的形状，先验概率代表所有点分到这个簇类的概率

哈哈一开始每一个中心点的先验概率都是1/k，协方差矩阵为单位矩阵，均值就是中心点

EM算法分为两个主要步骤首先是期望步骤，用算法评估每个数据点x属于每个聚类C的概率 $P(C|x)$ 。这个步骤不像是kmeans中直接通过最短距离把x分类到最近的簇中，而是先计算出x分类到所有簇的概率

之后是最大化步骤Maximum，用前一个步骤也就是期望步骤算出来的概率来更新每个簇类的参数

最后使得来自刚刚E步骤的下的似然对数函数最大

■ Expectation Step

$$p(C_j | x_i) = \frac{p(x_i | C_j) \cdot p(C_j)}{p(x_i)} = \frac{p(x_i | C_j) \cdot p(C_j)}{\sum_j p(x_i | C_j) \cdot p(C_j)}$$

■ Maximization Step

$$\mu_j = \frac{\sum_i p(C_j | x_i) \cdot x_i}{\sum_i p(C_j | x_i)} \quad \Sigma_j = \frac{\sum_i p(C_j | x_i) \cdot (x_i - \mu_j) \cdot (x_i - \mu_j)^T}{\sum_i p(C_j | x_i)} \quad p(C_j) = \frac{\sum_i p(C_j | x_i)}{N}$$

好，那第一个算exp的，首先p(C)初始化为1/k，就带进去算，算出来tm一堆 $p(C_j|x_i)$ ，好，现在有了每一个点属于每一个簇的期望值（可能性）

之后重新校准每一个簇的属性包括均值和协方差

重新计算他们的平均值和协方差，用极大似然算出每一个簇的均值和方差参数

同时更新这几个簇的先验概率