1.题目

**GMM练习**

2.数据集

自行生成、或者下载。

3.任务描述

（1）GMM极大似然估计公式推导（自己用纸笔完成，拍照，以图片形式嵌入到ipynb文件中(Markdown Cell)，图片和ipynb文件打包提交）；

（2）1维/2维GMM self-implement示例代码分析与运行效果（分析提供的示例代码，增加注释的Cell）；

（3）self-implement GMM vs. scikit learn GMM（比对运行效率，比对聚类效果，包括可视化方式，以及Within cluster sum of squares）。

（4）K-Means vs. GMM比较（比对不同数据集的聚类效果，包括可视化方式，以及Within cluster sum of squares）

备注：自行选用参考文献[1]的数据集。

备注：Within cluster sum of squares：The sum of the squared deviations from each observation and the cluster centroid.

4.成果提交

以ipynb形式提交。

5.参考文献

[1] https://archive.ics.uci.edu/ml/datasets.php