

软件环境：

MATLAB_R2014B

内容：

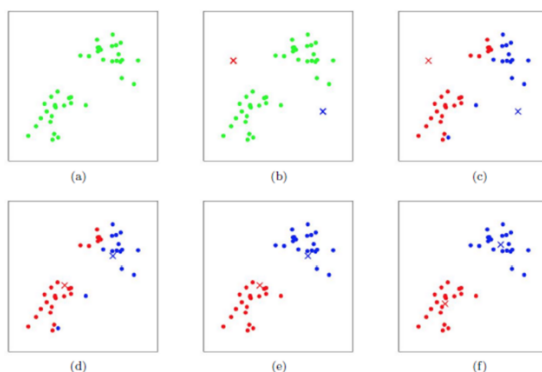
用 MATLAB 实现 k-means 算法，并对给定数据进行聚类

步骤：

一. 理解 k-means 算法

1. 随机在范围内取 k 个点,
2. 把所有数据分别计算和 k 个点距离，按距离分到离自己最近的点
3. 取 k 个类的中心点
4. 重复第二部第三步，直到这几个中心点的变动范围在一个很小的范围内

下图展示了对 n 个样本点进行K-means聚类的效果，这里 k 取2。



二. 用 matlab 实现

1. 把数据2d-data.mat， 3d-data.mat放到workspace中
2. 通过下面的代码把数据写入，然后用 r 就是用来这个 500×2 的矩阵了

```
load('2d-data.mat');
```

3. 在所有数据中随机取 k 个数

```
% 取所有点里面的k个  
point = [];  
num=[];  
num=fix(rand(1,k)*col);  
% 下面这个矩阵要转一下  
point=data(num,:).'
```

4. 通过 for 循环计算点与点之间的欧式距离，选择点的分类

5. 取每个类的所有点的 x, y, z 。位置的平均值，得到新的中心位置
6. 计算各类中心点与初始点的差值，如果不在设置范围内则不跳出循环，若在范围外则跳出循环：

```
while(err>0.01)
```

7. 最后运行以后得到结果，然后显示出来

