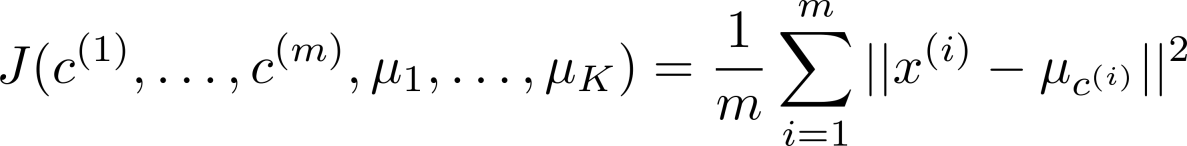
### KMean

1. 算法分为两步骤:
2. 第一个for循环，计算每个样本数据x(i)距离哪一个聚类中心近，既:对于每个样本数据x(i),计算其应该属于的聚类中心，用c(i)(i=1,....m)来存储与第i个样本数据x(i)最近的聚类中心索引
3. 第二个for训练，是聚类中心的移动，即：对于每一个聚类中心，重新计算它的中心**μ**(k)
4. 代价



计算样本与地标点的距离

第一个循环用于减少c(i)引起的代价

第二个循环则是用于减小μk引起的代价

迭代的过程一定会是每一次迭代都在减小代价函数，不然便是出现了错误。

1. 通常用‘肘部法则’来确定类别个数k的值

4.K-Means主要有两个最重大的缺陷：

* 1. K 需事先给定
  2. K-Means算法需要随机初始种子点，这个随机种子点太重要，不同的随机种子点会有得到完全不同的结果