**第一个公式是k-means的评价函数，后面再看懂迭代就够了。**

k-means 所要优化的目标函数：设我们一共有 N 个数据点需要分为 K 个 cluster ，k-means 要做的就是最小化



rnk这个函数，其中在数据点 n 被归类到 cluster k 的时候为 1 ，否则为 0 。直接寻找和来最小化并不容易，不过我们可以采取迭代的办法：先固定，选择最优的，很容易看出，只要将数据点归类到离他最近的那个中心就能保证最小。下一步则固定，再求最优的。将对求导并令导数等于零，很容易得到最小的时候应该满足：

=mean(xn),其中xn为属于cluster k的点的坐标亦即的值应当是所有 cluster k 中的数据点的平均值。由于每一次迭代都是取到的最小值，因此只会不断地减小（或者不变），而不会增加，这保证了k-means 最终会到达一个极小值。虽然 k-means 并不能保证总是能得到全局最优解，但是对于这样的问题，像 k-means 这种复杂度的算法，这样的结果已经是很不错的了。

为什么说每一次迭代 都在减小？ 假设第 次迭代的聚类结果为，然后重新计算第个类别的中心为，

.

假设在本次被分到第2类，即,，但在下一次迭代被分到第1类()，即，其余点的类别不变。重新计算第个类别的中心为，则