

三、问题三的模型与建立

3.1 攀岩难度评级

3.1.1 YDS 系统介绍

本文引用 YDS 作为评价攀岩难度的模型系统，1937 年 Welzenbach 的攀登难度系统由山峦俱乐部引入美国，之后衍生为优胜美地十进制系统(YDS)，它将难度分为 6 个大类。

- (1) 徒步
- (2) 几乎用不到手的爬行
- (3) 需要携带绳子，但是几乎用不到的爬行
- (4) 通常需要绳子，很容易找到天然的保护点，脱落是致命的
- (5) 技术攀登，要有一定的攀岩技术，绳子和保护是必须的
- (6) 必须借助器械才能进行的攀登

其中这里的第五个级别就是指攀岩，也是常说的五级技术攀登，到了 20 世纪 50 年代，五级攀登由于需要被划分成 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 5.10, 前面的数字 5 仍然代表五级攀登，后面的 0 到 10 代表攀登路线的难易程度，这是个封闭的系统，最简单的线路被定义为 5.0，而最难的线路被定为 5.10。

3.1.2 评价指标选择

通过查阅相关资料，本文将选取路径总长度以及岩石之间的最大距离差，所攀爬岩石的个数作为评价指标，以评价该模型，并采用 YDS 系统的评分方式，以下是问题二路线的指标统计结果。

表 3.1 指标统计结果

终点编号	路径总长度	岩石之间最大距离差	攀爬岩石个数	总减分
111	1205.84	97.7394(7 → 15)	17	31
112	1254.05	97.7394(7 → 15)	17	33
113	1216.8	97.7394(7 → 15)	17	33
114	1277.48	99.0202(28 → 33)	18	31

3.2 Kmeans 聚类分析

3.2.1 Kmeans 算法原理

已知数据集 (x_1, x_2, \dots, x_n) ，Kmeans 聚类要把这 n 个数据划分到 k 个集合中 $(k \leq n)$ ，使得组内平方和最小，它的目标是找到使得下式满足的聚类 S_i

$$\arg \min_S \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (4-1)$$

其中 μ_i 是 S_i 中所有点的均值。

3.2.2 Kmeans 算法步骤

- (1) 对数据集进行标准化和归一化，避免均值和方差大的数据对聚类产生决定性影响
- (2) 选择初始化的 k 个样本作为初始聚类中心 $a = a_1, a_2 \dots a_k$
- (3) 针对数据集中每个样本 x_i ，计算它到 k 个聚类中心的距离，并将其分到距离

最小的聚类中心所对应的类中

(4) 针对每个类别 $a_j(j = 1,2..k)$ ，重新计算它的聚类中心 $a_j = \sum_{x \in S_i} x$ ，即属于该类的所有样本的质心

(5) 重复(3)(4)步，知道达到某个中止条件（迭代次数，可允许最小误差等）

其伪代码为：

获取数据 n 个 m 维的数据

随机生成 K 个 m 维的点

while(t)

for(int i=0;i < n;i++)

for(int j=0;j < k;j++)

计算点 i 到类 j 的距离

for(int i=0;i < k;i++)

1. 找出所有属于自己这一类的所有数据点

2. 把自己的坐标修改为这些数据点的中心点坐标

end

3.2.3 Kmeans 模型求解

根据机器学习中的肘部准则，选取斜率最大的 $K = 2$ 的分类个数肘部图，确定了以聚类为 2 能更好地描述评价指标的组合。

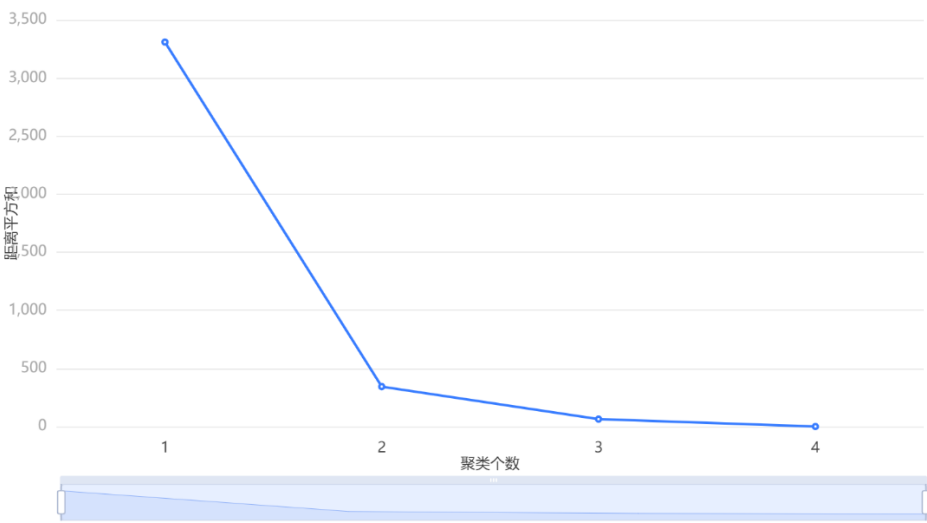


图 3.1 聚类数对比图

表 3.2 数据集聚类标注表

聚类种类	终点编号	路径总长度	岩石之间最大距离差	攀爬岩石个数	总减分
1	111	1205.84	97.7394	17	31
2	112	1254.05	97.7394	17	33
1	113	1216.8	97.7394	17	33
2	114	1277.48	99.0202	18	31

通过 Python 的可视化库 matplotlib 可以直观地展现出聚类散点图，由于变量数大于 2 个，即取主成分分析(PCA)降维后前两个主成分来绘制散点图，在一定程度上可查看聚类效果。

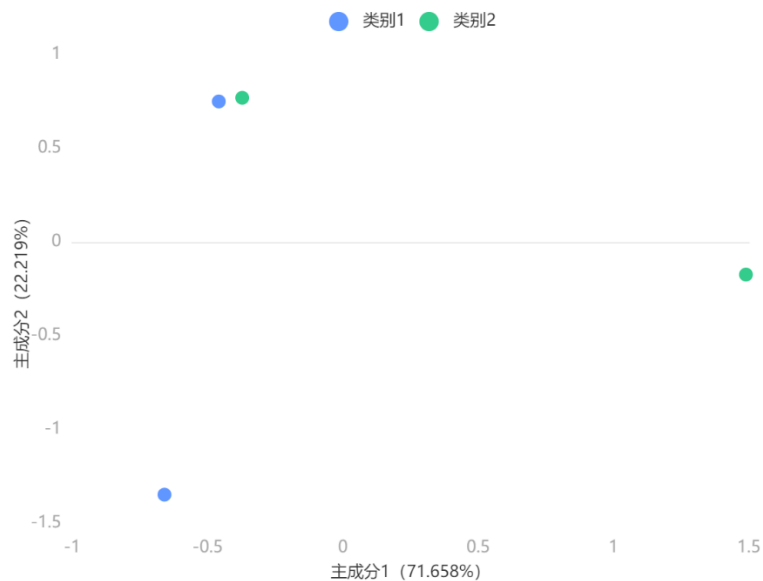


图 3.2 聚类散点图

由于是通过 Dijkstra 算法选择的最多得分路径，且与最短路径基本吻合，在聚类分析分为了两类的情况下，可将这 4 条路线分为 2 条评价等级为优的路线和 2 条评价等级为良的路线，经过映射可得到问题三的最终结果表如下所示：

表 3.3 问题三路线评级

终点编号	路径	评价等级(YDS)
111	1 → 7 → 15 → 24 → 29 → 34 → 40 → 48 → 57 → 60 → 69 → 75 → 84 → 91 → 100 → 103 → 111	5.1
112	1 → 7 → 15 → 24 → 29 → 34 → 40 → 48 → 57 → 60 → 69 → 78 → 88 → 93 → 101 → 107 → 112	5.3
113	1 → 7 → 15 → 24 → 29 → 34 → 40 → 48 → 57 → 60 → 69 → 78 → 88 → 93 → 101 → 110 → 113	5.2
114	1 → 7 → 15 → 24 → 29 → 34 → 40 → 48 → 57 → 60 → 69 → 78 → 88 → 89 → 94 → 96 → 105 → 114	5.4