

基于多元统计法对古玻璃文物的成分分析及鉴定

摘 要

本文主要研究了对古玻璃制品成分的分析 and 鉴别的问题。古玻璃制品随着时间的推移其表面难免会发生风化，而风化程度在外部上与其类型、纹饰、颜色有关，在内部与各化学成分有着千丝万缕的联系。同时，根据风化程度我们可以对玻璃进行一定的分类鉴定，这对玻璃后续的实际用途有指导性意义。

针对问题一，本文首先对玻璃文物的表面风化与其类型、纹饰、颜色进行**斯皮尔曼相关性分析**，得到玻璃的表面风化与其所属类型的相关性较强，与其纹饰、颜色有一定关系但不强烈；接着本文将古玻璃分为高钾和铅钡两类，每一类玻璃都对其表面是否风化与各化学成分进行了斯皮尔曼相关性分析，得到每一类玻璃与其是否风化相关性强的化学成分；为了方便描述风化与化学成分含量的统计规律，本文引入了正负相关和风化速率两概念，建立了预测模型与**单目标规划模型**，根据已知风化点检测数据，预测了其风化前的化学成分含量。

针对问题二，本文先对附件数据进行**数据可视化**，归纳总结出了高钾玻璃和铅钡玻璃的分类规律；对于两类别的玻璃，我们沿用问题一中每类玻璃的是否风化与各化学成分含量的相关性结论选择了合适的化学成分，并且利用**K-means 聚类**对每类玻璃进行了亚类划分。结果是高钾玻璃里又分了两个亚类，铅钡玻璃又分了两个大类。最后本文通过提取每一亚类最显著的化学成分的含量特征来对其合理性进行说明，以及对其进行敏感性分析，发现结果有其合理性，敏感性较低。

针对问题三，题目要求我们对附件表单 3 中的未知类别的玻璃文物进行化学成分分析进而鉴别其所属类型即建立判别标准；我们通过 **Logistics 回归**、**相关性优化**和**K-means 聚类**三个模型分别建立了对应的判别阈值 p_{cre} 建立了完善判别体系，并通过三个模型的对比得到了全变量 **Logistics** 的预测准确率最高的结论；对于灵敏度分析我们以 5%为度量进行改变，通过对照前后预测概率 p_i 的变化量得到了模型灵敏度高的结论。

针对问题四，玻璃类别样品按问题二中的亚类分类进行划分，同时我们通过**Pearson 相关性分析**对各亚类化学成分之间的关联关系进行探究并利用对照分析其差异性得到了风化亚类之间仅发生关联系数大小的改变；类别亚类还发生了主要影响化学成分改变的结论。

关键词：*Spearman*相关性分析 单目标规划 *Logistics* 回归 *K-means* 聚类

一、问题重述

1.1 问题背景

汉代的丝绸之路是中西文化交流的重要途径，在早期，玻璃是最有价值的实物证据之一。而我国古代的玻璃，是在吸取了西方的制造技术后，在国内就地取材制作而成的。所以，我国古代玻璃制品与国外的玻璃产品外形类似，但是其化学组成不同。

众所周知，玻璃的主要原材料为硅砂，其化学成份为二氧化硅（ SiO_2 ）。因为石英砂熔点很高，在精炼过程中，必须加入助熔剂，以减少熔融温度。根据加入助溶剂的不同，可以将古代玻璃制品分为铅钡玻璃和高钾玻璃。其中，铅钡玻璃是以铅矿石作为助熔剂烧制而成的，其铅、钡、氧含量较高；而高钾玻璃是以草木灰作为助熔剂烧制而成的，其钾含量较高；此外，古玻璃在埋地条件下，很容易受到侵蚀。在风化作用下，由于其内部和周围的物质之间存在着较多的相互转换，使其组成比例发生了改变，从而影响了对其分类的正确判断。

现有一批我国古代玻璃制品的相关数据，根据这些文物样品的化学成分，考古工作者使用检测手段已将其分为高钾玻璃和铅钡玻璃两种类型。具体的相关数据见附件。

1.2 数据说明

（1）附件中的表单 1 提供了关于古玻璃的基本信息，即文物的编号、纹饰、类型、颜色和是否表面风化。

（2）附件中的表单 2 提供了已分类古玻璃的化学成分比例。其中，文物取样点是在该文物的表面上随意取样的，其风化特征与附件中的表单 1 所示的文物相符；严重风化点属于风化层内的一点，未风化点属于风化玻璃表面未风化区域的一点；同一文物的部位 1 和部位 2 属于文物不同的两个部位，它们的成分和含量可能有差异。

（3）附件中的表单 3 提供了还没有分类的古玻璃的化学组分比例。

1.3 问题提出

（1）根据附件中表单 1 的数据建立数学模型，探究分析古玻璃的表面风化与其类型、纹饰、颜色的关系；根据已分类的玻璃的类型，建立数学模型探究出风化后和风化前化学成分含量的统计规律，并基于这个统计规律，根据风化后风化点的检测数据，预测其风化前的化学含量。

（2）根据附件数据建立数学模型探究玻璃分为高钾玻璃和铅钡玻璃的分类规律；对于每一类玻璃，选择合适的化学成分和方法再对其进行亚类划分，并对

划分的结果进行合理性和敏感性的分析。

(3) 建立数学模型探究分析附件表单 3 中还没有分类的古玻璃的化学成分，鉴别它们是所属铅钡玻璃还是高钾玻璃，并对分类结果的敏感性进行分析。

(4) 对于不同类别的古玻璃样品建立数学模型分析它们化学成分间的关联关系，并比较这种关联关系的差异性。

二、模型假设

1. 附件所提供的数据全部真实、准确、可靠。
2. 假设附件表单 3 的玻璃一定是铅钡或高钾类型。
3. 由于检测手段等原因，数据中可以存在古代玻璃文物的化学成分比例的累加和非 100% 的情况。
4. 将古代玻璃文物的化学成分比例的累加和介于 85%~105% 之间的数据视为有效数据。

三、符号说明

| 符号 | 文字说明 |
|--------------------------|-----------------------|
| ρ | 斯皮尔曼相关系数 |
| y_i | 需预测的风化样本化学成分含量 |
| x_i^{start}, x_i^{end} | 同一玻璃样本第 i 个化学成分的含量 |
| U_i | 需预测样本对于第 i 个化学成分的含量 |
| ε | 残差项 |
| x_{ij} | 标准化后的数据 |
| s_j | 均方差 |
| p_i | 铅钡类型玻璃的预测概率 |
| γ | 皮尔逊相关系数 |

四、问题分析

4.1 问题一的分析

问题一可以分为两个部分，第一个部分要求我们分析古玻璃文物的表面风化与其类型、纹饰和颜色的关系，第二个部分要求我们根据不同的玻璃类型，探讨文物表面风化与其化学成分含量的统计规律，并根据这个统计规律预测古玻璃风化前的化学成分含量。针对第一部分，即对玻璃文物的风化与其类型、纹饰、颜

色做一个相关性分析，归纳分析出风化与哪些指标的相关性更强烈，对第二题的分类提供依据；针对第二部分，根据题干，需先将玻璃分为高钾和铅钡两类进行分开讨论，然后沿用第一部分的相关性分析找到每一类玻璃的风化程度对哪些化学成分含量影响程度大。为了验证其合理性，利用 *EXCEL* 的原始数据绘制直方图使数据可视化，直观验证相关性分析的合理性。接着，为了方便描述统计规律，我们引入了正负相关和风化速率两个概念来描述统计规律。并在此基础上建立预测函数和单目标规划模型，求出古玻璃风化前化学成分含量的预测值。

4.2 问题二的分析

问题二可以分为两个部分，第一部分要求我们基于附件数据分析玻璃分为高钾、铅钡的分类规律，第二部分要求我们对每个类别选择合适的化学成分建立模型进行亚类划分，并对分类结果的合理性和敏感性分析。针对第一部分，我们先对附件中的数据进行了预处理，再利用 *EXCEL* 软件计算出四组数据中剩余 12 种化学元素的平均所占比例，并绘出图表进行数据可视化，在此基础上，归纳总结出分类规律；针对第二部分，首先将古玻璃文物分为高钾和铅钡两种情况讨论，再根据问题一相关性的结论选择合适的化学成分利用 *K-Means* 聚类进行亚类划分，最后进行合理性和敏感性分析。

4.3 问题三的分析

针对问题三，要想对附件表单 3 中的未知类别的玻璃文物进行化学成分分析进而鉴别其所属类型即建立判别标准；我们通过 *Logistics* 回归、相关性优化和 *K-means* 聚类三个模型分别建立了对应的判别阈值建立了完善判别体系，并通过三个模型的对比得到预测准确率最高的模型；对于灵敏度分析我们打算以 5% 为度量进行改变，通过对照前后预测概率 p_i 的变化量得到模型灵敏度。

4.4 问题四的分析

针对问题四，要想分析不同类别的玻璃化学成分之间的关系，我们需要先对玻璃类别样品按问题二中的亚类分类进行划分，同时我们通过 *Pearson* 相关性分析对各亚类化学成分之间的关联关系进行探究并利用对照分析得到其差异性。

五、模型的建立及求解

5.1 问题一模型的建立及求解

5.1.1 数据预处理

由题干和本题假设可以得知，本题将古代玻璃文物的化学成分比例的累加和介于 85%~105%之间的数据视为有效数据。因此对于附件中的表单 2，我们利用 EXCEL 表格的 SUM 函数将各文物的化学成分比例相加，发现编号为 15、17 号的文物化学成分比例相加不在有效范围内，故剔除该两组数据。

而针对附件中表单 1 的数据，由于检测手段等原因，部分文物颜色数据缺失。又由于颜色数据属于定性的分类数据，若进行插值补齐，还要考虑其可能作为一种新颜色新分类的数据而存在。并且，文物颜色数据的有无与考古学家的检测手段有关，其缺失属于完全随机缺失。因此我们选择剔除缺失值样本。一共剔除的数据如下表：

表 1：附件中剔除的数据

| 编号 | 类型 | 颜色 | 化学成分含量和 | 删除原因 |
|----|----|----|---------|--------------|
| 15 | 高钾 | 深绿 | 79.47% | 成分含量和不在有效范围内 |
| 17 | 高钾 | 浅蓝 | 71.89% | 成分含量和不在有效范围内 |
| 19 | 铅钡 | | 97.25% | 数据缺失 |
| 40 | 铅钡 | | 98.57% | 数据缺失 |
| 48 | 铅钡 | | 99.17% | 数据缺失 |
| 58 | 铅钡 | | 98.76% | 数据缺失 |

5.1.2 第一部分：玻璃文物指标的相关性分析

注意到附件表单 1 的数据属于定性、离散的分类数据，因此我们采用适用范围广的斯皮尔曼相关系数对玻璃文物的表面风化与其类型、纹饰和颜色的相关性进行分析，为了进一步显化内在规律，我们需要对离散变量进行赋值。

a.对于离散数据的赋值

针对古玻璃文物的四个指标：纹饰、类型、颜色、是否表面风化，我们进行变量赋值。例如纹饰这个指标，所有玻璃文物只有 A、B、C 三种离散变量，三种离散变量之间没有大小关系，为无序多项变量，通常情况下为了便于量化分析，我们将其分别赋值为 1、2、3，它们之间是相互平等、相互独立的关系。其他指标的离散变量具体赋值如下表：

表 2：离散变量的赋值

| | | | | | |
|------|----|----|-----|---|----|
| 纹饰 | 变量 | A | B | C | |
| | 赋值 | 1 | 2 | 3 | |
| 类型 | 变量 | 高钾 | 铅钡 | | |
| | 赋值 | 1 | 2 | | |
| 颜色 | 变量 | 蓝绿 | 浅蓝 | 紫 | 深绿 |
| | 赋值 | 1 | 2 | 3 | 4 |
| | 变量 | 浅绿 | 黑 | 绿 | 深蓝 |
| | 赋值 | 5 | 6 | 7 | 8 |
| 是否风化 | 变量 | 风化 | 无风化 | | |
| | 赋值 | 1 | 0 | | |

b. 基于斯皮尔曼 (Spearman) 相关系数模型的建立

对于大数据类题目，斯皮尔曼相关系数主要适用于两列变量及具有等级变量性质具有线性关系的数据，能够很好地处理序列中的相同值和异常值。其等效于应用于第 a 列 x_a 和第 b 列 y_b 的秩序的皮尔逊相关系数。对于样本容量为 n 的样本， n 个原始数据被转化成等级数据，相关系数 ρ 为：

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \quad (1)$$

而原始数据依据其在总体数据中平均的降序位置，被分配了一个相应的等级。

但是，实际应用中，变量间的连结是无关紧要的，于是可以通过简单的步骤计算 ρ 。对于第 a 列 x_a 和第 b 列 y_b 的相关性，设 a, b 列的相关系数为 $\rho(a, b)$ ，为：

$$\rho(a, b) = 1 - \frac{6 \sum d^2}{m(m^2 - 1)} \quad (2)$$

其中， d 为两个列的秩之差，即 a, b 列的等级差； m 为每列的长度。 ρ 值的值域为 $[-1, 1]$ ，相关系数的绝对值越大，关联度越高，即相关系数越接近于 1 或-1，两者关联度越高，相关系数越接近于 0，两者关联度越低。

c. 基于斯皮尔曼 (Spearman) 相关系数模型的求解

将附件中表单 1 指标中的离散变量赋值后，利用 MATLAB 中的内置函数求解，得到斯皮尔曼相关系数的有关热力图（具体代码见附录），如下图所示：



图 1：古玻璃文物指标间的相关性分析

d.结果分析

由古玻璃文物指标间的相关性分析图，因为 $\rho(\text{是否风化}, \text{类型}) = 0.4741$ ，故古玻璃文物的表面风化与其所属的类型有较强的正相关性；而 $\rho(\text{是否风化}, \text{纹饰}) = 0.2356$ ，可以得知古玻璃文物的表面风化与其纹饰有一定正相关性，但不是很强；而 $\rho(\text{是否风化}, \text{颜色}) = -0.1144$ ，可以得知古玻璃文物的表面风化与其颜色有一定负相关性，但不是很强。

5.1.3 第二部分：探究文物表面有无风化化学成分含量的统计规律

题干要求结合玻璃的类型，分析文物表面有无风化化学成分含量的统计规律。于是我们根据古玻璃文物的类型，我们将其分为高钾玻璃和铅钡玻璃两种类型。

a.古玻璃文物的风化程度与化学成分含量的相关性分析

对于古玻璃文物，我们首先需要知道玻璃文物在风化前后，哪些化学成分含量受到的影响最大，即风化与哪些化学成分含量的相关性较大。这里，我们仍然沿用第一部分斯皮尔曼相关系数的方法，分别得到高钾玻璃的风化程度与化学成分含量的 spearman 相关性系数，如下表所示：

表 3：高钾玻璃的风化与各化学成分含量的斯皮尔曼相关性

| | | | | | | | |
|------|---------|---------|---------|---------|---------|---------|---------|
| 风化程度 | 二氧化硅 | 氧化钠 | 氧化钾 | 氧化钙 | 氧化镁 | 氧化铝 | 氧化铁 |
| | 0.8178 | -0.3147 | -0.7056 | -0.5455 | -0.6016 | -0.795 | -0.5455 |
| 风化程度 | 氧化铜 | 氧化铅 | 氧化钡 | 五氧化二磷 | 氧化锶 | 氧化锡 | 二氧化硫 |
| | -0.2727 | -0.5426 | -0.3743 | -0.6023 | -0.487 | -0.1715 | -0.3147 |

(1)针对高钾玻璃的风化程度与各化学成分含量，由于 $\rho(\text{风化}, \text{二氧化硅}) = 0.8178$ ，故高钾玻璃的风化程度与二氧化硅的含量具有强烈的正相关性。同理，我们也可以得到高钾玻璃的风化程度与氧化钾、氧化钙、氧化镁、氧化铝、氧化铁、氧化铅、五氧化二磷具有较强的负相关性；高钾玻璃的风化程度与氧化钠、氧化铜、氧化钡、氧化锶、氧化锡、二氧化硫有一定负相关性，但不是很强。

表 4：铅钡玻璃的风化与各化学成分含量的斯皮尔曼相关性

| | | | | | | | |
|------|--------|--------|---------|--------|--------|---------|---------|
| 风化程度 | 二氧化硅 | 氧化钠 | 氧化钾 | 氧化钙 | 氧化镁 | 氧化铝 | 氧化铁 |
| | -0.835 | -0.373 | -0.1988 | 0.4067 | -0.052 | -0.347 | -0.0204 |
| 风化程度 | 氧化铜 | 氧化铅 | 氧化钡 | 五氧化二磷 | 氧化锶 | 氧化锡 | 二氧化硫 |
| | 0.3168 | 0.6938 | 0.1171 | 0.5833 | 0.3679 | -0.0865 | 0.2917 |

(2)针对铅钡玻璃的风化程度与各化学成分含量，我们可以得到铅钡玻璃的风化程度与二氧化硅、氧化铜的含量具有强烈的负相关性；铅钡玻璃的风化程度与氧化钙、五氧化二磷、二氧化硫的含量具有强烈的负相关性，与其他化学成分有一定相关性，但不够强烈。

b.基于 EXCEL 表中原始数据绘制的直方图对相关性分析的检验

对于上文古玻璃文物与化学成分含量的斯皮尔曼相关性分析，我们利用 EXCEL 表中的原始数据绘制直方图来进行检验。

(1) 针对高钾玻璃的风化程度与各化学成分相关性分析的检验，我们根据表中高钾玻璃的数据，绘制了风化与未风化下高钾玻璃各化学成分含量的变化，如下直方图所示：

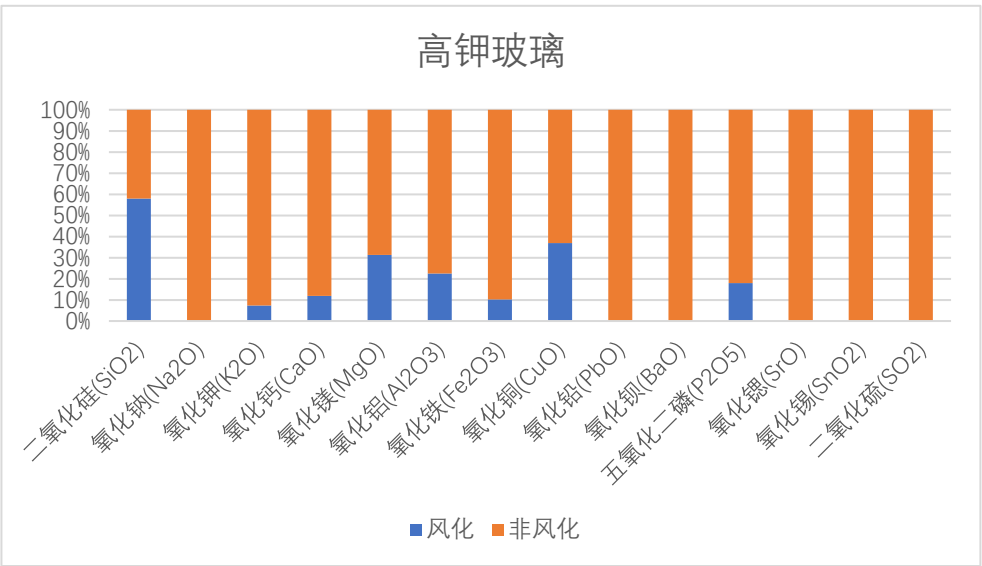


图 2：风化与未风化下高钾玻璃各化学成分含量的变化

分析上图，可以验证高钾玻璃的风化程度与二氧化硅的含量具有强烈的正相关性。同理，我们也可以验证高钾玻璃的风化程度与氧化钾、氧化钙、氧化镁、氧化铝、氧化铁、氧化铅、五氧化二磷具有较强的负相关性。

(2) 针对铅钡玻璃的风化程度与各化学成分相关性分析的检验，我们根据表中铅钡玻璃的数据，选取并绘制了 49 号文物样本风化与未风化下铅钡玻璃各化学成分含量的变化，和 8 号文物样本风化与严重风化下铅钡玻璃各化学成分含量的变化，如下直方图所示：

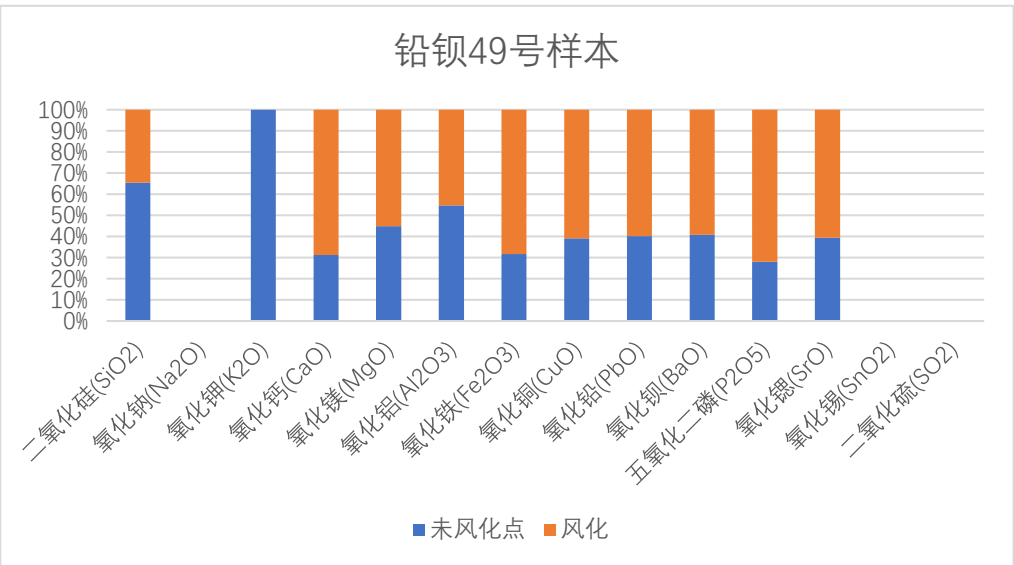


图 3：风化与未风化下 49 号铅钡玻璃各化学成分含量的变化

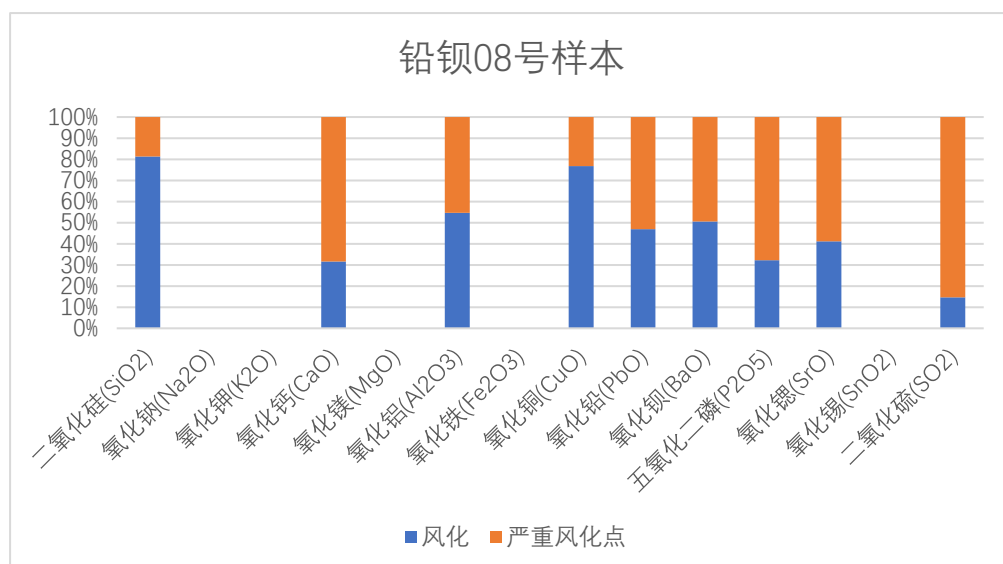


图 4：风化与严重风化下 08 号铅钡玻璃各化学成分含量的变化

分析上两张图，可以验证铅钡玻璃的风化程度与二氧化硅、氧化铜的含量具有强烈的负相关性；铅钡玻璃的风化程度与氧化钙、五氧化二磷、二氧化硫的含量具有强烈的负相关性，与其他化学成分有一定相关性，但不够强烈。

c.探究分析文物样品表面风化程度与化学成分含量的统计规律

在上文找到了文物样品表面的风化程度与哪些化学成分含量关系强烈后，为了方便得到其统计规律，我们引入了正负相关和风化速率两个概念来描述统计规律。

其中，正负相关在相关性分析中有所体现。若相关系数为正，则对应的风化程度与其化学成分的量变化为正相关；若相关系数为负，则对应的风化程度与其化学成分的量变化为负相关。

对于风化速率，我们可以用古玻璃样品各化学成分在风化前后的变化来进行描述。针对高钾玻璃，利用 MATLAB 绘制出高钾玻璃各化学成分变化平均值图，如下：

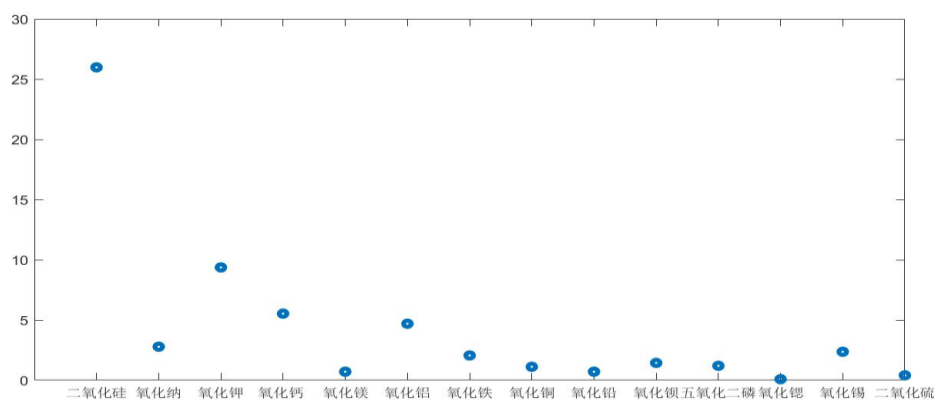


图 5：高钾玻璃各化学成分变化平均值

针对铅钡玻璃，我们选取了 8 号和 26 号铅钡玻璃样本并用 MATLAB 绘制出铅钡玻璃各化学成分变化图来比较铅钡玻璃风化与未风化下其各化学成分含量变化如下图所示：

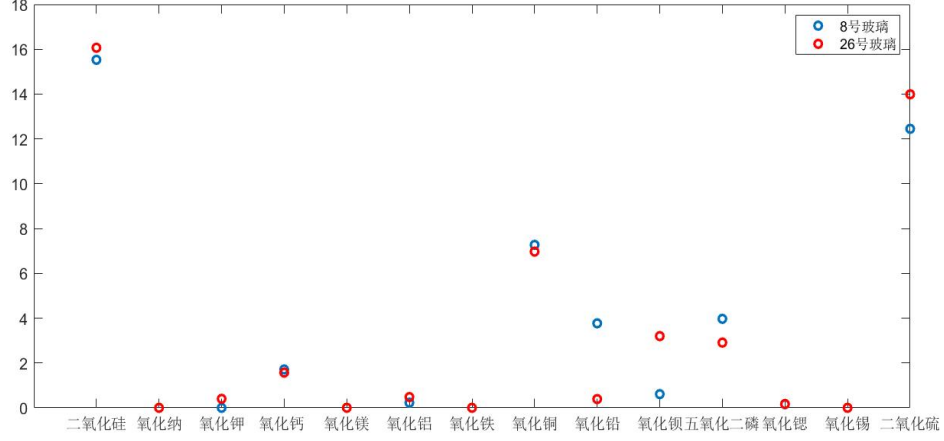


图 6: 8 号和 26 号铅钡玻璃的各化学成分变化量

根据上图，综上，我们以可视化的角度得知了古玻璃的风化程度与玻璃化学成分的含量具有显著的统计规律，这对下文我们定量建立预测函数和建立单目标规划模型提供依据。

d. 预测函数的建立和单目标规模模型的建立

于上文我们已经得知了风化与玻璃化学成分的含量具有显著的统计规律；基于此，我们可以通过分析同一件玻璃样品的风化点与未风化点之间的化学成分差异，以及需预测的风化样本化学成分含量 y_i 建立预测函数 $f_i(k)$ 如下：

$$f_i(k) = \frac{x_i^{strat}}{x_i^{end}} \quad i \in (1, 2 \cdots 14) \quad (3)$$

其中 x_i^{strat} 、 x_i^{end} 表示同一玻璃样本第 i 个化学成分的含量；则对于每一个需要预测的化学成分含量指标都能由预测函数确定一个预测系数，同时上文确定的预测系数与需预测的风化样本化学成分含量 y_i 满足如下关系：

$$U_i = y_i \cdot f_i(k) \quad i \in (1, 2 \cdots 14) \quad (4)$$

其中 U_i 表示需预测样本对于第 i 个化学成分的含量。同时考虑到因检测手段等原因可能导致其成分比例的累加和非 100% 的情况，引入检测损失误差 ε 来描述其他方面的误差：

$$U_i = y_i \cdot f_i(k) + \varepsilon \quad i \in (1, 2 \dots 14) \quad (5)$$

由于本题中若成分比例的累加和位于 85% ~ 105% 之间，则视数据为有效数据，较未风化样本累加和应为 100% 的成分性高出 5%，因此可初步定义检测损失误差的取值区间如下：

$$\varepsilon \in (-0.05, 0.05) \quad (6)$$

同时根据题意可知未风化样本各成分累加和应趋近 100%；基于此，我们可以建立关于预测比重函数 U_i 的一个单目标规划模型：

$$\min(\sum_{i=1}^{14} U_i(\varepsilon) - 100\%) \quad (7)$$

于上文我们建立了以未风化样本各成分累加和趋近 100% 的单目标规划模型，并通过最优的规化预测函数对风化样本进行未风化时化学成分的含量预测，其中目标函数如下：

$$\min(\sum_{i=1}^{14} U_i(\varepsilon) - 100\%) \quad (8)$$

由上文各符号的定义可知，决策变量 ε 及待预测变量 y_i 等均具有取值范围限制；因此为了确保模型的合理性，给出各变量范围限制如下：

① 已知样本风化前后的化学成分含量取值范围：

$$0 \leq x_i^{strat}, x_i^{end} \leq 1 \quad (9)$$

② 预测比重函数的化学成分取值范围：

$$0 \leq U_i \leq 1 \quad (10)$$

③ 合理的检测损失误差的取值范围：

$$-0.05 < \varepsilon < 0.05 \quad (11)$$

综上我们建立了单目标规划模型如下：

$$\min(\sum_{i=1}^{14} U_i(\varepsilon) - 100\%) \quad (12)$$

$$s.t. \begin{cases} f_i(k) = \frac{x_i^{strat}}{x_i^{end}} & i \in (1, 2 \dots 14) \\ U_i = y_i \cdot f_i(k) + \varepsilon & i \in (1, 2 \dots 14) \\ \min(\sum_{i=1}^{14} U_i(\varepsilon) - 100\%) \\ 0 \leq x_i^{strat}, x_i^{end} \leq 1 \\ 0 \leq U_i \leq 1 \\ -0.05 < \varepsilon < 0.05 \end{cases} \quad (13)$$

e. 单目标规模模型的求解

求出预测的古玻璃文物风化前的各化学成分的含量。由于篇幅原因，展示了部分结果（具体结果见附录），如下：

表 5：部分预测的古玻璃文物风化前的各化学成分的百分含量

| 编号 | 08 铅钡 | | 02 铅钡 | | 10 高钾 | | 09 高钾 | |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 风化后 | 预测值 | 风化后 | 预测值 | 风化后 | 预测值 | 风化后 | 预测值 |
| SiO_2 | 20.14 | 42.90 | 36.28 | 77.28 | 96.77 | 70.01 | 95.02 | 68.75 |
| Na_2O | 0 | 0 | 0 | 0 | 0 | 2.7 | 0 | 2.7 |
| K_2O | 0 | 0 | 1.05 | 0.315 | 0.92 | 11.49 | 0.59 | 7.37 |
| CaO | 1.48 | 0.99 | 2.34 | 1.57 | 0.21 | 1.54 | 0.62 | 4.56 |
| MgO | 0 | 0 | 1.18 | 1.06 | 0 | 0 | 0 | 0 |
| Al_2O_3 | 1.34 | 1.97 | 5.73 | 8.43 | 0.81 | 2.78 | 1.32 | 4.53 |
| Fe_2O_3 | 0 | 0 | 1.86 | 0.77 | 0.26 | 2.27 | 0.32 | 2.80 |
| CuO | 10.41 | 6.54 | 0.26 | 0.16 | 0.84 | 1.44 | 1.55 | 2.66 |
| PbO | 28.68 | 19.67 | 47.43 | 32.54 | 0 | 0.7 | 0 | 0.7 |
| BaO | 31.23 | 16.02 | 0 | 0 | 0 | 1.436 | 0 | 1.436 |
| P_2O_5 | 3.59 | 2.19 | 3.57 | 2.18 | 0 | 0 | 0.35 | 1.60 |
| SrO | 0.37 | 0.18 | 0.19 | 0.09 | 0 | 0.083 | 0 | 0.083 |
| SnO_2 | 0 | 0 | 0 | 0 | 0 | 2.36 | 0 | 2.36 |
| SO_2 | 2.58 | 0.38 | 0 | 0 | 0 | 0.407 | 0 | 0.406 |

f. 预测结果误差分析

于上文我们已经建立了由风化化学成分数据预测未风化前化学成分的预测函数，为了进一步讨论模型的合理性选取附件表单 2 中的第 50 号样本的风化点与未风化点通过比较预测值与真实值的大小进行误差分析如下：

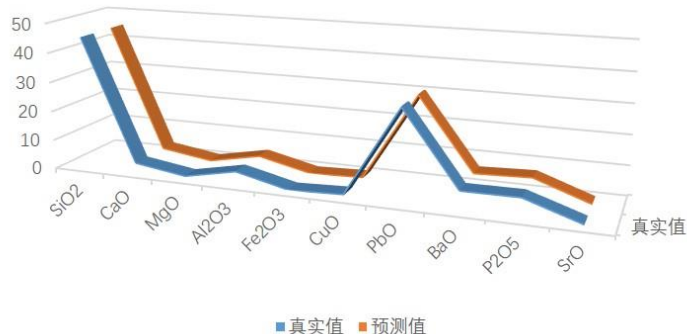


图 7：各化学成分真实值与预测值的对比

同时绘制真实值与预测值的差值折线图如下：

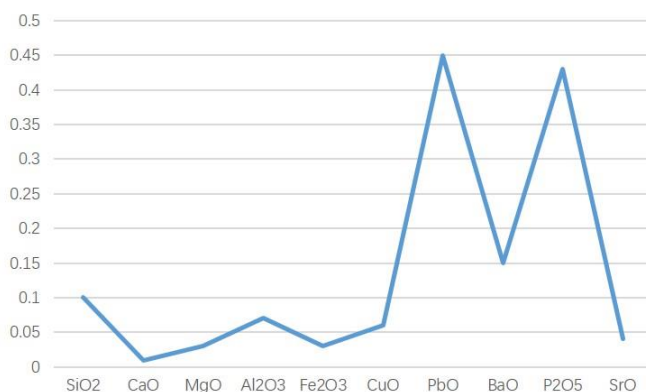


图 8：各化学成分真实值与预测值差值变化

由上图可知通过我们建立的预测函数对第 50 号样本得到的预测值与真实值之间的误差始终低于 0.5，且两者具有相同的变化规律；因此进一步说明了我们模型在本题中的合理性。

5.2 问题二模型的建立及求解

5.2.1 第一部分：基于数据可视化归纳总结出不同类别玻璃的分类规律

a.数据预处理

观察附件表单 2 的数据，发现 69 次检测当中只有 6 个玻璃样本内含有氧化锡(SnO_2)，8 个玻璃样品内含有二氧化硫(SO_2)，很难从较少的数据中得出统计学结果，故下文在讨论高钾玻璃、铅钡玻璃的分类规律时不再考虑这两个元素的作用效果。

即使是同一块玻璃，在检测不同风化程度的位置时，化学元素的差异也很大。所以我们将高钾玻璃和铅钡玻璃再划分为风化、无风化共四组。

b.数据可视化

在 EXCEL 中计算出四组数据中剩余 12 种化学元素的平均所占比例，由于篇幅原因，论文只展示部分结果如下表所示：

表 6：四组玻璃各化学成分所占百分比

| 类型 | 有无风化 | SiO_2 | Na_2O | K_2O | ... | P_2O_5 | SrO |
|----|------|---------|---------|--------|-----|----------|-------|
| 高钾 | 风化 | 93.96 | 0 | 0.82 | ... | 0 | 0 |
| | 未风化 | 67.98 | 2.78 | 10.17 | ... | 1.53 | 0.08 |
| 铅钡 | 风化 | 38.8 | 2.59 | 0.25 | ... | 5.49 | 0.47 |
| | 未风化 | 24.91 | 1.41 | 0.39 | ... | 5.72 | 0.47 |

为了更加直观的看出两种类型玻璃的差异，绘制如下折线图：

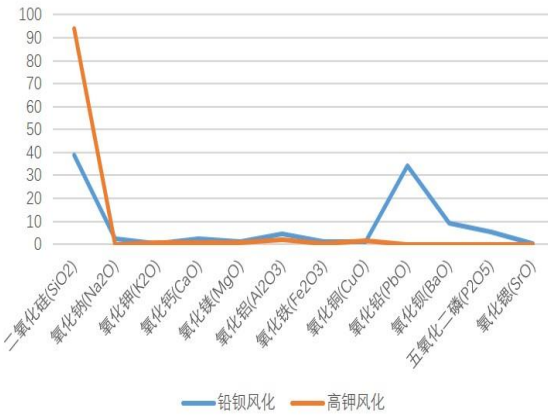


图 9：风化的玻璃元素占比对比图

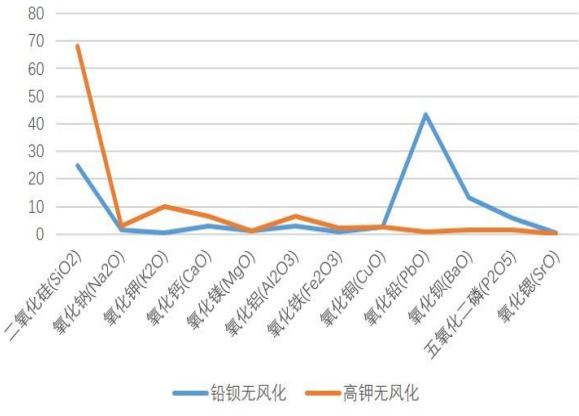


图 10：未风化的玻璃元素占比对比图

c. 结果分析与分类规律

（1）观察图 1 中的数据，发现在风化情况下，相较于铅钡玻璃，钾玻璃的二氧化硅含所占百分比要高很多，铅钡玻璃的氧化钡所占百分比比高钾玻璃高很多，此外，氧化铝、氧化钡、五氧化二磷这三种含量相对较少的元素，在铅钡玻璃中所占百分比都高于高钾玻璃。

（2）观察图 2 中数据，发现在无风化情况下，高钾玻璃的二氧化硅含量要原告于铅钡玻璃，同样，铅钡玻璃的氧化铅含量远高于高钾玻璃。此外，风化后的高钾玻璃的氧化钾、氧化铁、氧化钙这几种元素所占百分比上升，超过了铅钡玻璃。

综上所述，我们得到分类规律：在此基础上，风化前的高钾玻璃氧化钾、氧化铁、氧化钙所占百分比要高于铅钡玻璃，风化后铅钡玻璃的氧化铝所占百分比要高于高钾玻璃；高钾玻璃内的二氧化硅所占百分比靠近 75%，古铅钡玻璃的氧化铅含量位于 20-45%之间，而二氧化硅的含量位于 30-50%之间。

5.2.2 第二部分：基于 K-Means 聚类对不同类别玻璃进行亚类划分

题干要求对于每一类玻璃，选择合适的化学成分和方法再对其进行亚类划分，并对划分的结果进行合理性和敏感性的分析。同样地，首先将古玻璃文物分为高钾和铅钡两种情况讨论。接着沿用第一问相关性的结论进行不同类别玻璃化学成分的选取。最后建立 K-Means 聚类模型对不同类别玻璃进行亚类划分，并对合理性和敏感性分析。

a. 不同类别玻璃化学成分的选取

(1) 针对高钾玻璃，沿用第一问高钾玻璃的风化程度与其各化学成分的相关性结论，选择相关性较大的化学成分，如下表所示：

表 7：高钾玻璃中与风化程度相关性较大的化学成分

| | | | | |
|----------------------|--------|---------|---------|---------|
| | 二氧化硅 | 氧化钾 | 氧化钙 | 氧化镁 |
| $\rho(\text{与风化程度})$ | 0.8178 | -0.7056 | -0.5455 | -0.6016 |
| | 氧化铝 | 氧化铁 | 氧化铅 | 五氧化二磷 |
| $\rho(\text{与风化程度})$ | -0.795 | -0.5455 | -0.5426 | -0.6023 |

(2) 针对铅钡玻璃，同理，沿用第一问铅钡玻璃的风化程度与其各化学成分的相关性结论，选择相关性较大的化学成分，如下表所示：

表 8：铅钡玻璃中与风化程度相关性较大的化学成分

| | | | | | |
|----------------------|--------|--------|--------|--------|--------|
| | 二氧化硅 | 氧化铜 | 氧化钙 | 五氧化二磷 | 二氧化硫 |
| $\rho(\text{与风化程度})$ | -0.835 | 0.3168 | 0.4067 | 0.5833 | 0.2917 |

b. K-Means 聚类分析模型的建立

K-means 聚类算法是一种迭代求解的聚类分析算法，其步骤是随机选取 k 个对象作为初始的聚类中心，然后计算每个对象与各个种子聚类中心之间的距离，如果新计算出来的质心和原来的质心之间的距离小于某一个设置的阈值，我们可以认为聚类已经达到期望的结果，算法终止。如果新质心和原质心距离变化很大，那么我们再重新确定聚类中心，再次每个对象与各个种子聚类中心之间的距离，这样经历 3-5 次的循环，最终确定聚类中心以及各类所包含的点。

c. K-Means 聚类分析模型的求解

对于古玻璃文物，利用 K-Means 聚类对其进行亚类划分，即划分成 k 类，并给出每个样本数据对应的中心点，具体算法步骤如下：

Step1 为了消除各指标之间的单位不同，以及数值量级之间差别的影响，我们

第一步需要对数据进行标准化处理，对 i 个被评价对象的 j 项指标的指

标值 x_{ij} 有标准化处理:

$$x'_{ij} = \frac{x_{ij} - \bar{x}}{s_j} \quad (14)$$

其中 $s_j = \sqrt{\frac{1}{n} \sum_1^n (x_{ij} - \bar{x})^2}$ ($i \neq j$) 为被评价指标关于第 i 个指标的均方差。

Step2 随机选取 k 个中心, 记为 $u_1^{(0)}, u_2^{(0)}, \dots, u_k^{(0)}$ 。

Step3 定义损失函数即各个样本距离所属簇中心点的误差平方和:

$$J(c, \mu) = \min \sum_{i=1}^M \|x_i - u_{c_i}\|^2 \quad (15)$$

Step4 令 $t = 0, 1, 2, \dots$ 为迭代函数, 重复如下过程知道 J 收敛:

Step(4.1) 对于每一个样本 x_i , 将其分配到距离最近的中心:

$$c_i^t < -\arg \min_k \|x_i - u_k^t\|^2 \quad (16)$$

Step(4.2) 对于每一类中心 k , 重新计算该类的中心

$$\mu_k^{t+1} < -\arg \min_u \sum_{i: c_i^t = k}^b \|x_i - u\|^2 \quad (17)$$

根据上述算法流程, 利用 MATLAB 求解得到聚类数和聚类中心点坐标。

(1) 针对高钾玻璃, 结果如下表所示:

表 9: 高钾玻璃相关化学成分的聚类数和各化学成分含量中心值

| 种类 | SiO_2 | K_2O | CaO | MgO | Al_2O_3 | Fe_2O_3 | PbO | P_2O_5 |
|----|---------|--------|-------|-------|-----------|-----------|-------|----------|
| 1 | 91.29 | 2.234 | 0.904 | 0.375 | 2.336 | 0.1987 | 0.125 | 0.463 |
| 2 | 64.93 | 10.68 | 6.198 | 0.900 | 7.233 | 2.317 | 0.600 | 1.481 |

根据上表中的聚类数及各化学成分中心值, 可以得到高钾玻璃的亚类划分, 在此基础上, 我们又展现了每一亚类高钾玻璃最显著的成分含量均值, 划分结果如下表:

表 10: 高钾玻璃亚类划分结果图

| 种类 | 高钾玻璃编号 |
|----|-------------------------|
| 1 | 03、07、09、10、12、18、22、27 |
| 2 | 01、04、05、06、13、14、16、21 |

(2) 针对铅钡玻璃, 结果如下表所示:

表 11: 铅钡玻璃相关化学成分的聚类数和各化学成分含量中心值

| 种类 | SiO_2 | CaO | CuO | P_2O_5 | SO_2 |
|----|---------|-------|-------|----------|--------|
| 1 | 25.34 | 2.64 | 2.50 | 4.95 | 1.22 |
| 2 | 58.50 | 1.20 | 0.98 | 0.90 | 0.18 |

根据上表中的聚类数及各化学成分中心值，可以得到铅钡玻璃的亚类划分，划分结果如下表：

表 12: 铅钡玻璃亚类划分结果图

| 种类 | 铅钡玻璃编号 |
|----|--|
| 1 | 02、08、11、19、20、24、26、30、34、36、 38、39、40、41、43、51、52、54、56、57、58 |
| 2 | 23、25、28、29、31、32、33、35、37、42、 44、45、46、47、48、49、50、53、55 |

5.2.3 对分类结果的合理性分析

(1) 针对高钾玻璃，利用 EXCEL 分别对亚类里的玻璃中的化学成分含量进行归纳分析，寻找每个亚类里的玻璃最显著的共同特征，并计算出最显著的化学成分的含量均值，归纳如下表：

表 13: 高钾玻璃亚类最显著成分的含量均值

| 种类 | 高钾玻璃编号 | 最显著成分的含量均值 |
|----|-----------------------------|----------------------------------|
| 1 | 03、07、09、10、12、18、 22、27 | $SiO_2(91.29\%)$ |
| 2 | 01、04、05、06、13、14、 16、21 | $K_2O(10.68\%)$, $CaO(6.198\%)$ |

分析分类结果，我们不难发现高钾玻璃中的第一类最显著的成分是二氧化硅。查阅相关资料可知，虽然玻璃中最主要的成分就是石英砂（二氧化硅），但一般的玻璃二氧化硅的含量只有 71%~73%，因此像这种二氧化硅纯度极高的高钾玻璃可以视为石英玻璃，其具有很强的光学、热学性质，可以用于制作高温灯具等；

(2) 针对高钾玻璃，利用 EXCEL 分别对亚类里的玻璃中的化学成分含量进行归纳分析，寻找每个亚类里的玻璃最显著的共同特征，并计算出最显著的化学成分的含量均值，归纳如下表：

表 14: 铅钡玻璃亚类最显著成分的含量均值

| 种类 | 铅钡玻璃编号 | 最显著成分的含量均值 |
|----|--------|------------|
|----|--------|------------|

| | | |
|---|--|-----------------------------|
| 1 | 02、08、11、19、20、24、 26、30、34、36、38、39、 40、41、43、51、52、54、 56、57、58 | $P_2O_5(4.95\%)$ |
| 2 | 23、25、28、29、31、32、 33、35、37、42、44、45、 46、47、48、49、50、53、 55 | $PbO(12.68\%), BaO(8.89\%)$ |

分析结果，我们不难发现铅钡玻璃中第一类最显著的成分是五氧化二磷，查阅相关资料可知，其可视为磷酸盐玻璃，可以作为激光增益介质，对稀土离子有高兼容性，可用于制作光通信放大器等；

同样地，我们发现铅钡玻璃中的第二类最显著的成分是氧化铅和氧化钡，作为常规的铅钡玻璃，其质软，透明性好，易加工，可用于制作高级器皿和装饰品。

综上，我们发现使用 K-Means 聚类分析划分的铅钡玻璃的每个亚类都在其化学组成含量上有显著特征，并与我们查阅的相关资料相符。此次分类对铅钡玻璃的实际用途有指导意义，可以认为第二部分使用的 K-Means 聚类分析有合理性。

5.2.4 对分类结果的敏感性分析

上文我们已经通过选择每种玻璃影响显著的化学成分作为分类指标，利用 *K-means* 聚类分析得出了附件表单的亚类共四种，同时题目进一步要求对模型进行灵敏度分析。

对于 *K-means* 聚类，我们以 5% 为度量改变其影响显著的化学成分含量再同上文利用 *K-means* 观察前后的聚类结果改变；同时对照改变前后的聚类结果，通过两者之间的变换来描述模型的灵敏度；以 5% 为度量改变其显著的化学成分聚类结果对照如下表：

表 15：高钾玻璃初聚类和再聚类结果对比图

| 种类 | 初聚类高钾玻璃编号 | 再聚类高钾玻璃编号 |
|----|-----------------------------|--------------------------------|
| 1 | 03、07、09、10、12、18、 22、27 | 03、07、09、10、12、22、 27 |
| 2 | 01、04、05、06、13、14、 16、21 | 01、04、05、06、13、14、 16、18、21 |

表 16: 铅钡玻璃初聚类 and 再聚类结果对比图

| 种类 | 初聚类铅钡玻璃编号 | 再聚类铅钡玻璃编号 |
|----|--------------------|--------------------|
| 1 | 02、08、11、19、20、24、 | 02、08、11、19、20、24、 |
| | 26、30、34、36、38、39、 | 26、30、34、37、38、39、 |
| | 40、41、43、51、52、54、 | 40、41、43、51、52、54、 |
| | 56、57、58 | 56、57、58 |
| 2 | 23、25、28、29、31、32、 | 23、25、28、29、31、32、 |
| | 33、35、37、42、44、45、 | 33、35、36、42、44、45、 |
| | 46、47、48、49、50、53、 | 46、47、48、49、50、53、 |
| | 55 | 55 |

通过上表我们可以得知：对于 *K-means* 聚类分类结果以 5% 为度量的影响是不显著的，改变化学成分含量后的聚类结果仍然与原始结果保持高度一致。由此可见对于本题，我们的 *K-means* 聚类分析模型具有良好的稳定性，进一步说明了模型的合理性。

5.3 问题三模型的建立及求解

问题三要求我们对附件表单 3 中未知类别的玻璃文物依据其化学成分进行分析鉴别玻璃文物的所属类型即：铅钡玻璃、高钾玻璃；通过观察表单 3 数据可知原题可转化为已知玻璃的各化学成分含量进行玻璃的类型预测。

由于其结果为 0、1 分类变量，其决策变量为连续型变量，因此我们可以通过对附件表单 2 中的原始数据进行 *Logistics* 回归并且确定判断阈值 p_{cre} ，从而建立满足题述要求的一种量化分类模型。

5.3.1 基于 *Logistics* 回归分类模型的建立

Logistics 回归是一种研究分类结果与决策变量之间关系的多变量分析方法，对于本文由于玻璃文物是否风化会对其化学成分造成显著性的影响，因此原始数据需按是否风化分为两大类分别进行回归以确保模型的合理性。

以风化的玻璃为例，若记分类变量铅钡玻璃为 1，高钾玻璃为 0 且玻璃文物是铅钡类型的预测概率为 p_i 则可以建立 *Logistics* 回归方程如下：

$$\ln \frac{p_i}{1-p_i} = b_0 + b_1 x_{1i} + b_2 x_{2i} + \cdots + b_{13} x_{13i} + b_{14} x_{14i} + \varepsilon \quad i \in (1, 2 \cdots 24) \quad (18)$$

其中， b_0 为函数截距， $b_1, b_2 \cdots b_{14}$ 是与 $x_{1i}, x_{2i} \cdots x_{14i}$ 无关的回归系数， ε 是引入的随机误差项。

同时为了方便讨论预测概率对 p_i 进行显化有：

$$p_i = \frac{1}{1 + e^{-(b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_{13} x_{13i} + b_{14} x_{14i} + \varepsilon)}} \quad (19)$$

5.3.2 基于 Logistics 回归分类模型的求解

对原始数据进行合理的映射处理，就可以用同问题一的方法得到回归系数，从而求解得到回归方程；我们可以定义算法流程如下：

Step1 不妨定义任一总体 Y 为离散型的最大似然函数 $H(\alpha) = \prod_{i=1}^{34} p(x_{1i}, \alpha)$ ，并对

函数的两边同时取对数得到权重对数似然函数如下：

$$\ln H(\alpha') = \sum_{i=1}^{34} \ln p(x_{1i}, \alpha');$$

Step2 通过求得对数似然函数的各个权重偏导，其次令其累加和 0 即可得到最大似然估计值的最优解，同时由最大似然估计值可以确定一组最优相应的回归参数。

Step3 通过确立的最佳回归参数即可确定出上文建立的 *Logistics* 回归函数，基于此，定义附件表单 2 中的铅钡玻璃的最小与平均的中值预测概率作为附件表单 3 预测类别的阈值也即：

$$p_{cre} = \frac{p_{average} + \min(p_i)}{2} \quad (20)$$

Step4 于上文已经确定了 *Logistics* 回归函数以及通过附件表单 2 确定了预测阈值 p_{cre} ，再将附件表单 3 中的已风化的玻璃文物各化学成分含量作为自变量代入回归函数若大于阈值则为铅钡玻璃也即有判断函数如下：

$$P(i) = \begin{cases} 1 & , p_{pre}^i > p_{cre} \\ 0 & , p_{pre}^i < p_{cre} \end{cases} \quad (21)$$

按照上述算法流程分析求解，得到全变量下逻辑回归结果，如下表所示：

表 17：全变量下回归系数 b 的值

| b_0 | b_1 | b_2 | b_3 | b_4 | b_5 | b_6 | b_7 |
|-------|--------|--------|--------|-------|--------|--------|--------|
| 7.006 | -0.067 | -0.637 | -1.151 | 0.13 | -0.328 | -0.308 | -1.053 |

| | | | | | | |
|-------|-------|----------|----------|----------|----------|----------|
| b_8 | b_9 | b_{10} | b_{11} | b_{12} | b_{13} | b_{14} |
| 0.598 | 0.027 | -0.348 | 0.232 | -2.348 | 3.818 | 0.311 |

回归效果如下分类评价指标表所示：

表 18：分类评价指标表

| 准确率 | 召回率 | 精确率 | F1 | AUC |
|-------|-------|-------|-------|-------|
| 0.836 | 0.836 | 0.836 | 0.836 | 0.939 |

其中，残差项 $\varepsilon = 0$ 。令 $y^* = \ln \frac{p_i}{1-p_i}$ 则有全变量逻辑回归方程如下：

$$y^* = 7.006 - 0.067x_{1i} - 0.637x_{2i} - 1.151x_{3i} + 0.13x_{4i} - 0.328x_{5i} \\ - 0.308x_{6i} - 1.053x_{7i} + 0.598x_{8i} + 0.027x_{9i} - 0.348x_{10i} + 0.232x_{11i} \\ - 2.348x_{12i} + 3.818x_{13i} + 0.311x_{14i}$$

由分类评价指标表可知，此次回归的准确、召回、精确率较高，回归效果较好。但是，我们也应该考虑到实际情况中一些变量即一些化学成分含量对玻璃文物的影响并不显著，因此我们通过筛选主要变量重新建立 *Logistics* 回归模型进行优化。

5.3.3 利用相关性的 *Logistics* 回归优化模型的建立

其余相关性不强的变量造成的影响通过随机误差项 ε' 描述得到优化模型，同时研究角度从同一风化下的分类变量转化为同一类型下以是否风化作为 0、1 分类变量；

以铅钡玻璃为例，若记分类变量风化为 1，非风化为 0 且铅钡玻璃文物风化预测概率为 p_i' 则可以建立优化的 *Logistics* 回归方程如下：

$$\ln \frac{p_i'}{1-p_i'} = c_0 + c_1x_{1i} + c_2x_{4i} + c_3x_{8i} + c_4x_{11i} + c_5x_{14i} + \varepsilon' \quad i \in (1, 2 \cdots 24)$$

其中， c_0 为函数截距 $c_1, c_2 \cdots c_5$ 是与 $x_{1i}, x_{4i} \cdots x_{14i}$ 无关的回归系数， ε 是引入的随机误差项。同时为了方便讨论预测概率对 p_i' 进行显化有：

$$p_i' = \frac{1}{1 + e^{-(c_0 + c_1x_{1i} + c_2x_{4i} + c_3x_{8i} + c_4x_{11i} + c_5x_{14i} + \varepsilon')}}}$$

5.3.4 利用相关性的 *Logistics* 回归优化模型的求解

算法步骤同 5.3.2 全变量 *Logistic* 回归分类模型的求解，这里不多赘述。

按照上述算法流程分析求解，得到全变量下逻辑回归结果，如下表所示：

表 19：分变量下回归系数 b 的值

| b_0 | b_1 | b_2 | b_3 | b_4 | b_5 |
|-------|--------|--------|-------|--------|-------|
| 1.063 | -0.026 | -0.231 | 0.096 | -0.021 | 0.278 |

回归效果如下分类评价指标表所示：

表 20：分类评价指标表

| 准确率 | 召回率 | 精确率 | F1 | AUC |
|-------|-------|-------|-------|-------|
| 0.806 | 0.806 | 0.812 | 0.804 | 0.799 |

其中，残差项 $\varepsilon = 0$ 。令 $y^* = \ln \frac{p_i}{1-p_i}$ 则有求解得到的分变量逻辑回归方程：

$$y^* = 1.063 - 0.026x_{1i} - 0.231x_{2i} + 0.096x_{3i} - 0.021x_{4i} + 0.278x_{5i}$$

由分类评价指标表可知，此次回归的准确、召回、精确率虽较 5.3.2 中的模型略微降低，但其考虑了实际情况中一些变量即一些化学成分含量对玻璃文物的影响并不显著，且总体回归效果较好。

5.3.5 基于问题二的聚类分析模型的建立

由于附件表 3 的数据具有确定性即一定是属于风化铅钡、风化高钾、非风化铅钡、非风化高钾四类，因此利用 K-means 聚类将附件表单 3 的数据分成四类得到分类结果；再利用每一类的化学成分含量较聚类中点的大小比较，进一步确定每一类的具体物理意义即确定分类函数如下：

$$\Omega(i) = \begin{cases} \text{高钾} & , K > K_{average} \begin{cases} \text{风化} & , SiO_2 > SiO_{2average} \\ \text{非风化} & , SiO_2 < SiO_{2average} \end{cases} \\ \text{铅钡} & , SO_2 > SO_{2average} \begin{cases} \text{风化} & , SiO_2 < SiO_{2average} \\ \text{非风化} & , SiO_2 > SiO_{2average} \end{cases} \end{cases} \quad (22)$$

5.3.6 基于问题二的聚类分析模型的求解

算法步骤同 5.3.2 全变量 Logistic 回归分类模型的求解，这里不多赘述。根据上文通过 K-means 以 4 为聚类数将需要分类的附件表单 3 进行聚类可以得到结果如下表所示：

表 21：K-means 聚类结果

| 种类 | 玻璃编号 | 表面风化 | 表面未风化 |
|----|----------|-------|-------|
| 1 | A2、A3、A4 | A2 | A3、A4 |
| 2 | A1、A6、A7 | A6、A7 | A1 |
| 3 | A5 | A5 | |
| 4 | A8 | | A8 |

从上表可知聚类种类 1、2 中均同时存在风化和未风化样本的预测结果，由此可见 *K-means* 聚类虽然在已知数据的亚类划分上效果理想，但在玻璃类别预测的运用效果并不理想。

5.3.7 对结果的敏感性分析

(1) 对于全变量预测函数的敏感性分析

我们选择第 28 号玻璃样本进行敏感性分析，其各化学成分所占的百分比分别为：二氧化硅：68.08%、氧化钠：2.31%、氧化钾：0.26%、氧化钙：1.34%、氧化镁：1%、氧化铝：4.7%、氧化铁：0.41%、氧化铜：0.33%、氧化铅：17.14%、氧化钡：4.04%、五氧化二磷：1.04%、氧化锶：0.12%、氧化锡：0.23%、二氧化硫：1.96%。

令这些变量以 1%为步长，绘制在单一变量在上下 5%范围内波动时，因变量的变化情况。由于数据较多，为了结果更加清晰，此处选择了五种化学元素进行展示：

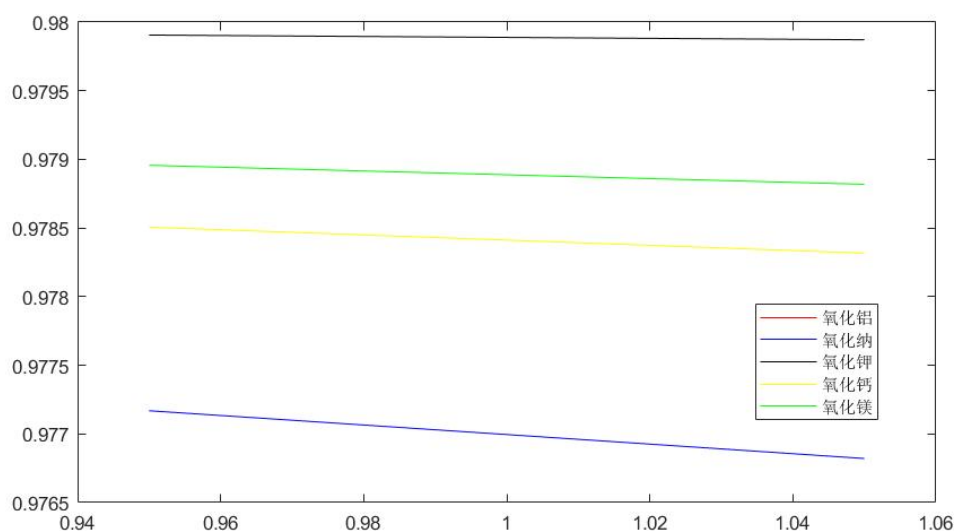


图 11: 全变量函数预测值随自变量波动的变化

从中我们可以看出，当自变量在 5% 的范围内发生波动时，整体函数的预测值变化不大，灵敏度较低，模型较为稳定。

(2) 以铅钡类型玻璃为例，本文选择 8 号铅钡玻璃进行灵敏度分析，其所含 5 个变量的初值分别为，二氧化硅：20.14%、氧化钙：1.48%、氧化铜：10.41%、氧化铅：28.68%、五氧化二磷 3.59%。

令这些变量以 1%为步长，上下波动 5%，利用 *MATLAB* 绘制在单一变量在 5%范围内波动时，因变量的变化情况，如下图所示：

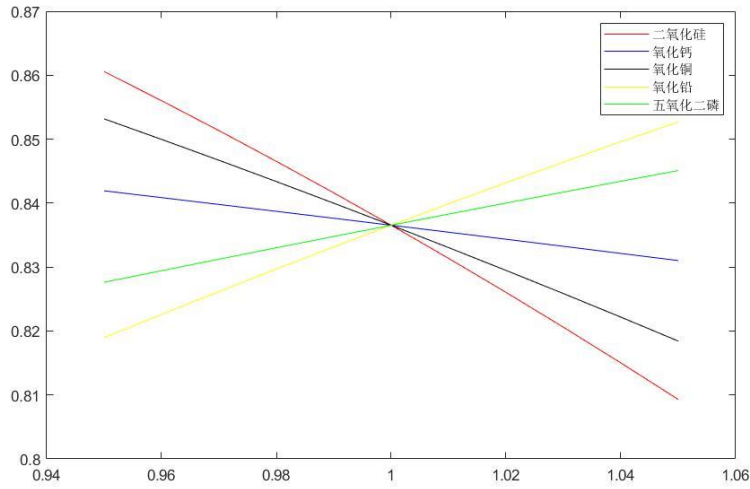


图 12: 8 号铅钡玻璃敏感性分析图

对结果预测概率随各变量变动时的变化。根据上图可以得知，预测概率与五个变量大致呈现一次函数的关系，虽然二氧化硅、氧化钙与预测概率构成的图像的斜率相对较高，但是总体上呈现低敏感性，由此可以确定模型时稳定的。

5.4 问题四模型的建立及求解

问题四要求我们对在上文问题二中分的不同亚类玻璃文物分析其化学成分之间的关联关系，同时比较不同亚类之间关联关系的差异性；考虑到其研究对象均为连续性变量，因此这里使用 *Pearson* 相关性分析更方便。

对于研究对象均为连续型变量的关联关系，我们通过 *Pearson* 相关性分析进行描述，同时通过分析系数矩阵得到各亚类化学成分之间的关系。

5.4.1 基于皮尔逊相关性分析模型的建立

于上文问题一中我们已经通过 *Sperman* 相关性分析得到了对于不同的玻璃类型的主要影响变量，为了进一步讨论不同亚类之间的关联关系，我们引入 *Pearson* 相关性分析定义如下：

Pearson 相关系数用于衡量定距变量间的线性关系，其数学表达如下：

$$\gamma = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N \sum x_i^2 - (\sum x_i)^2} \sqrt{N \sum y_i^2 - (\sum y_i)^2}} \quad (23)$$

式中， γ 表示 *Pearson* 相关系数； x 和 y 分别对应两组分析对象的数据序列也即本文的化学成分含量指标，同时可以得到各个化学成分之间的相关性系数矩阵以铅钡玻璃文物样品如下：

$$\begin{bmatrix} \gamma_{1-1} & \cdots & \gamma_{1-j} & \cdots & \gamma_{1-5} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \gamma_{i-1} & \cdots & \gamma_{i-j} & \cdots & \gamma_{i-5} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \gamma_{5-1} & \cdots & \gamma_{5-j} & \cdots & \gamma_{5-5} \end{bmatrix}$$

γ 值的值域为 $[-1,1]$ ，相关系数的绝对值越大，关联度越高，即相关系数越接近于 1 或-1，两者关联度越高，相关系数越接近于 0，两者关联度越低。

5. 4. 2 基于皮尔逊相关性分析模型的求解

根据上文的建立的相关性分析模型，通过 *SPSS* 数据分析得到四个亚类分别对应的相关性矩阵，并通过热力图进行可视化处理；基于此，根据相关性矩阵的系数总结各化学成分之间的关联关系。

(1) 对于未风化铅钡玻璃进行相关性分析得到系数矩阵如下：

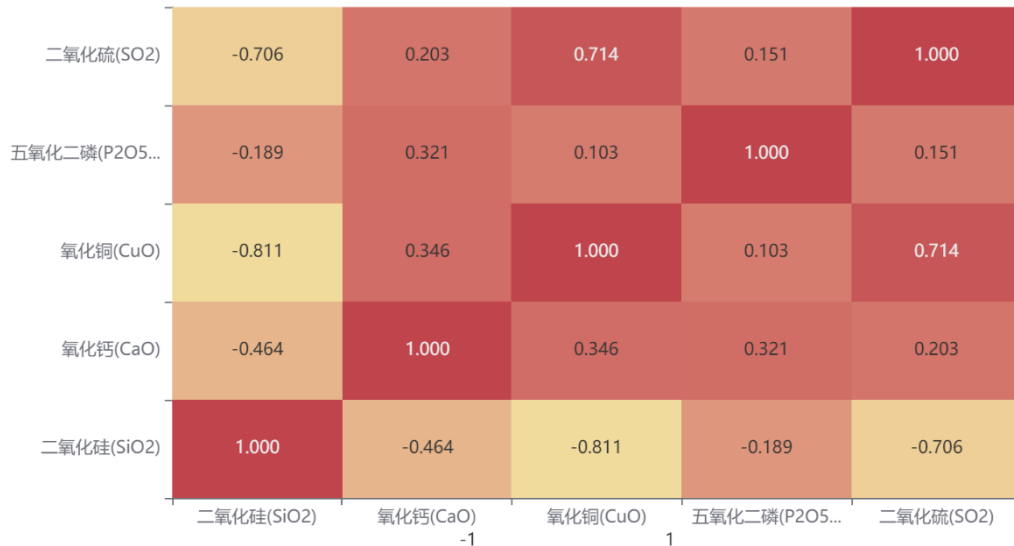


图 13：未风化铅钡玻璃的相关性系数矩阵

根据相关性系数矩阵我们可以得知：在非风化铅钡玻璃中，二氧化硅 SiO_2 与其余各成分呈负相关，且与氧化铜 CuO 的关联度最高，与五氧化二磷 P_2O_5 的关联性最低的变化规律。

(2) 同理，对于风化铅钡玻璃进行相关性分析得到系数矩阵如下：

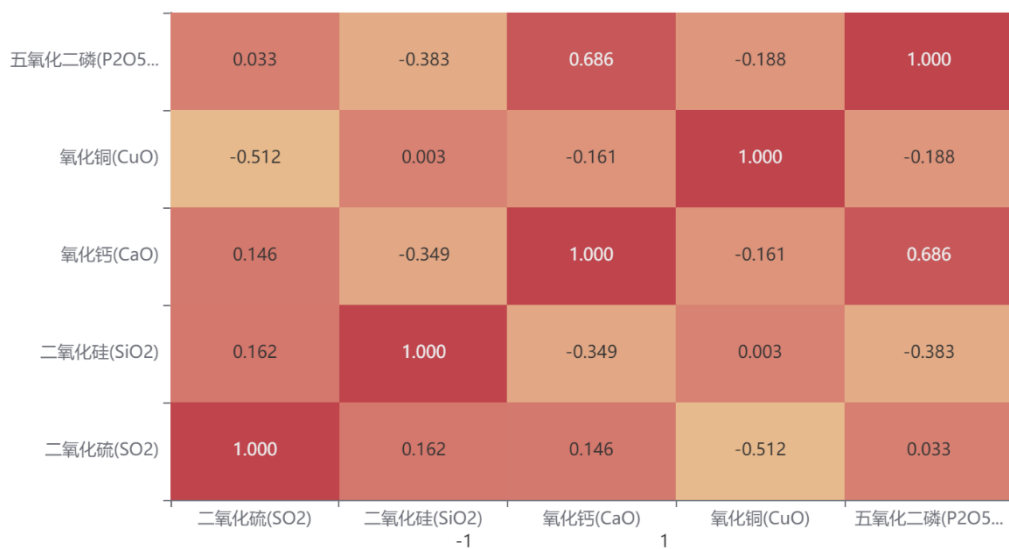


图 14: 风化铅钡玻璃的相关性系数矩阵

根据相关性系数矩阵我们可以得知：在风化铅钡玻璃中，二氧化硅 SiO_2 与氧化铜和二氧化硫 SO_2 呈正相关，与氧化钙 CaO 和五氧化二磷 P_2O_5 呈正相关，且与五氧化二磷的关联性最高的变化规律。

(3) 同理分别对于非风化和风化的高钾玻璃进行相关性分析得到系数矩阵如下：



图 15: 风化高钾玻璃的相关性系数矩阵

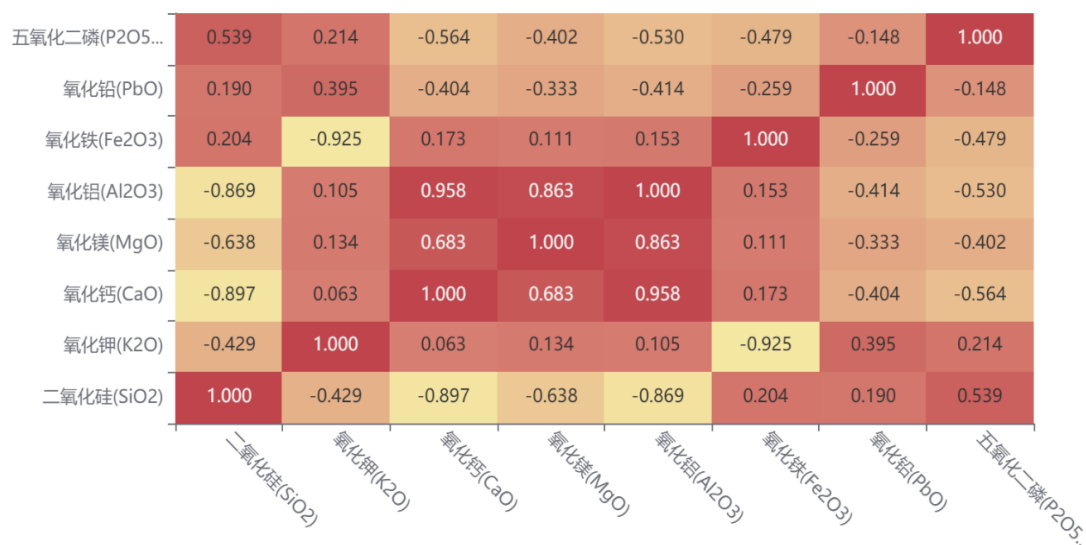


图 16: 未风化高钾玻璃的相关性系数矩阵

根据相关性系数矩阵我们可以得知：在非风化高钾玻璃中，二氧化硅 SiO_2 仅与氧化铁 Fe_2O_3 呈正相关，与其余变量均呈负相关且与氧化钾 K_2O 的关联度最高；在风化高钾玻璃中，二氧化硅 SiO_2 与 Fe_2O_3 、 PbO 、 P_2O_5 呈正相关，与其余变量均呈负相关且于氧化钙 CaO 的关联度最高。

5. 4. 3 对化学成分关联关系的差异性分析

综上所述我们已经通过Pearson相关性分析得到了各个亚类的相关系数矩阵并且提取出来了各化学成分之间的关联关系；根据题意需要进一步对亚类之间的关联关系进行讨论，为了方便讨论可将关联关系以表格形式给出如下：

表 22: 四类玻璃与已提取出的化学成分关联性

| | 相关性 | 化学成分 | 相关性最大的成分 |
|-------|-----|---|-----------|
| 高钾风化 | 正相关 | P_2O_5 、 PbO 、 Fe_2O_3 | P_2O_5 |
| | 负相关 | Al_2O_3 、 MgO 、 CaO 、 K_2O | CaO |
| 高钾未风化 | 正相关 | Fe_2O_3 | Fe_2O_3 |
| | 负相关 | P_2O_5 、 PbO 、 Al_2O_3 、 MgO 、 CaO 、 K_2O | K_2O |
| 铅钡风化 | 正相关 | SO_2 、 CuO | SO_2 |
| | 负相关 | CaO 、 P_2O_5 | P_2O_5 |
| 铅钡未分化 | 正相关 | 无 | 无 |
| | 负相关 | SO_2 、 CuO 、 CaO 、 P_2O_5 | CuO |

通过表格可知,对于亚类风化与未风化铅钡玻璃中二氧化硅 SiO_2 的变化规律有:两者的最高关联化学元素均为五氧化二磷 P_2O_5 ,且均呈正相关;但风化后的铅钡玻璃二氧化硅含量受化学成分较未风化的关联变化程度减小,即各成分之间的改变逐渐减小趋于稳定。

对于亚类风化与未风化高钾玻璃中二氧化硅 SiO_2 变化规律有:风化前高钾玻璃二氧化硅受氧化钾 K_2O 的负相关影响和氧化铁 Fe_2O_3 的正相关影响显著;而风化后受氧化钙 CaO 的负相关影响和五氧化二磷 P_2O_5 的正相关影响显著,也即风化前后不仅影响程度发生改变,同时主要影响的化学成分也发生了改变。

对于亚类铅钡玻璃和高钾中二氧化硅 SiO_2 变化规律有:风化前后铅钡玻璃中二氧化硅的主要影响的化学成分仅发生了关联程度变化,而高钾玻璃中不仅发生了关联程度变化而且主要影响化学成分也发生了改变。

六、模型的评价及推广

6.1 模型的优点

- 1.本文使用 EXCEL 软件对表格数据进行处理并做出各种图表,简便、直观、快捷。
- 2.运用多种数学软件和统计软件对数据进行处理和相关性分析,如 MATLAB, EXCEL, SPSS 等,取长补短,使计算结果更加准确。
- 3.本文将定量和定性方法相结合,对古玻璃文物风化程度与其化学成分的含量的关系进行了分析和总结。
- 4.本文建立的模型与实际紧密联系,利用聚类分析建立的数学模型成功对不同类别的玻璃进行再次的亚类划分,对玻璃的使用用途有直接的指导意义,有其合理性。

6.2 模型的缺点

- 1.模型中为使计算简便,使所得结果更加理想,忽略了一些次要影响的因素。
- 2.对于一些数据进行了必要的处理,会带来一些误差。
- 3.问题一中的预测模型若使用微分方程模型,预测结果则可能会更准确。

6.3 模型的推广

本文的应用背景是基于古代玻璃的化学成分的分析 and 鉴别,在相似背景下应用众多,如考古学中新出土的古陶瓷的化学成分分析和鉴别,亦或是地理学中风化对物质化学成分含量的影响等等;此外,问题二中的亚分类对玻璃材料的实际使用用途有直接的指导意义。

参考文献

- [1]王婕,李沫,马清林,等. 一件战国时期八棱柱状铅钡玻璃器的风化研究[J]. 玻璃与搪瓷,2014,42(2):6-13.
- [2]伏修锋,干福熹.基于多元统计分析方法对一批中国南方和西南地区的古玻璃成分的研究[J].文物保护与考古科学,2006(04):6-13.
- [3]郭富,金普军,李婷婷.四川汉源麦坪遗址出土陶器成分分析与研究[J].四川文物,2021(3):95-103
- [4]朱泓.黄河流域新石器时代居民体质特征的聚类分析[J]. 北方文物, 1990(4):5.
- [5]陈蕾蕾.多元统计分析在数学建模中的应用[J].无线互联科技,2020,17(24):98-99.

附录

附录 1

问题一 Spearman 相关性分析与图像绘制

```
%% spearman 相关性分析与热力图绘制

figure
% 求维度之间的相关系数
rho = corr(X11, 'type','kendall');
% 绘制热图
string_name={'纹饰','类型','是否风化','颜色'};
xvalues = string_name;
yvalues = string_name;
h = heatmap(xvalues,yvalues, rho, 'FontSize',12, 'FontName','宋体');
h.Title = '斯皮尔曼相关性';
colormap summer

%%类别分开
%%      各种元素之间的联系
K1_2(isnan(K1_2))=0;
Ba1_2(isnan(Ba1_2))=0;
figure    %铅钡玻璃
% 求维度之间的相关系数
rho = corr(Ba1_2, 'type','kendall');
% 绘制热图
string_name={'风化程度','二氧化硅','氧化钠','氧化钾','氧化钙 ','氧化镁
','氧化铝' ,'氧化铁' ,'氧化铜' ,'氧化铅' ,'氧化钡' ,'五氧化二磷' ,'氧化锶
','氧化锡' ,'二氧化硫'};
xvalues = string_name;
yvalues = string_name;
h = heatmap(xvalues,yvalues, rho, 'FontSize',12, 'FontName','宋体');
h.Title = '铅钡玻璃 kendall 相关性检验';
colormap summer

%%      高钾玻璃各元素之间 spearman 相关性分析与作图
```

```

figure %高钾玻璃
rho = corr(K1_2, 'type','spearman');
% 绘制热图
h = heatmap(xvalues,yvalues, rho, 'FontSize',12, 'FontName','宋体');
h.Title = '高钾玻璃斯皮尔曼相关性检验';
    colormap summer

%%    速率增长图
x=1:14;
V={'二氧化硅','氧化钠','氧化钾','氧化钙 ','氧化镁' ,'氧化铝' ,'氧化铁' ,'
氧化铜' ,'氧化铅' ,'氧化钡' ,'五氧化二磷' ,'氧化锶' ,'氧化锡' ,'二氧化硫
'};
%    08 号速率增长图
plot (x, PB,'o', 'linewidth', 5);
set(gca, 'xTick', [1:1:14]);
set(gca,'XTickLabel',V);
set(gca, 'FontSize', 12);
%    26 号速率增长图
plot (x,PB26, 'o', 'linewidth', 4);
set(gca, 'xTick', [1:1:14]);
set(gca,'XTickLabel',V);
set(gca, 'FontSize', 12);

%    49 号速率增长图
plot (x,PB49, 'o', 'linewidth', 4);
set(gca, 'xTick', [1:1:14]);
set(gca,'XTickLabel',V);
set(gca, 'FontSize', 12);

%    54 号速率增长图
plot (x,PB54, 'o', 'linewidth', 4);
set(gca, 'xTick', [1:1:14]);
set(gca,'XTickLabel',V);
set(gca, 'FontSize', 12);

```



```
% 高钾玻璃增长图
plot (x,Kave, 'o', 'linewidth', 4);
set(gca, 'xTick', [1:1:14]);
set(gca,'XTickLabel',V);
set(gca, 'FontSize', 12);
```

附录 2

第三题关于灵敏度检验

Function1 铅钡玻璃预测概率函数

```
function F=Bafun(x)
F=1/(1+exp(-(0.186*x(1)-0.539*x(2)-
0.244*x(3)+0.086*x(4)+0.356*x(5)+4.972))));
function F=Kfun(x)
```

Function2 高钾玻璃预测概率函数

```
F=1/(1+exp(-(1.216*x(1)+3.163*x(2)+11.412*x(3)+28.408*x(4)-
9.869*x(5)-17.35*x(6)+5.506*x(7)+23.612*x(8)...
+45.059*x(9)-144.702))));
```

Function2 整体预测概率函数、

```
function F=afun(x)
F=1/(1+exp(-(0.067*x(1)-0.637*x(2)-1.151*x(3)+0.13*x(4)-
0.328*x(5)-0.308*x(6)-1.053*x(7)+0.598*x(8)...
+0.027*x(9)-0.348*x(10)+0.232*x(11)-
2.348*x(12)+3.818*x(13)+0.311*x(14)+7.006))));
```

灵敏度绘图

```
%% 铅钡玻璃灵敏度分析
x8=[20.14 1.48 10.41 28.68 3.59];
X8=[x8;x8;x8;x8;x8;x8;x8;x8;x8;x8;x8];

%% 二氧化硅
X81=X8;
```

```

for i=1:11
    X81(i,1)=(0.94+0.01*i)*x8(1);
end
X81=X81';
y1=ones(11,1);
for i=1:11
    y1(i)=Bafun(X81(5*(i-1)+1:5*(i-1)+1+4));
end
%%    氧化钙
X82=X8;
for i=1:11
    X82(i,2)=(0.94+0.01*i)*x8(2);
end
X82=X82';
y2=ones(11,1);
for i=1:11
    y2(i)=Bafun(X82(5*(i-1)+1:5*(i-1)+1+4));
end
%%    氧化铜
X83=X8;
for i=1:11
    X83(i,3)=(0.94+0.01*i)*x8(3);
end
X83=X83';
y3=ones(11,1);
for i=1:11
    y3(i)=Bafun(X83(5*(i-1)+1:5*(i-1)+1+4));
end
%%    氧化铅
X84=X8;
for i=1:11
    X84(i,4)=(0.94+0.01*i)*x8(4);
end
X84=X84';

```

```

y4=ones(11,1);
for i=1:11
    y4(i)=Bafun(X84(5*(i-1)+1:5*(i-1)+1+4));
end
%% 五氧化二磷
X85=X8;
for i=1:11
    X85(i,5)=(0.94+0.01*i)*x8(5);
end
X85=X85';
y5=ones(11,1);
for i=1:11
    y5(i)=Bafun(X85(5*(i-1)+1:5*(i-1)+1+4));
end

%% 灵敏度分析图像
plot(0.95:0.01:1.05,y1,'r');
hold on
plot(0.95:0.01:1.05,y2,'b');
hold on
plot(0.95:0.01:1.05,y3,'k');
hold on
plot(0.95:0.01:1.05,y4,'y');
hold on
plot(0.95:0.01:1.05,y5,'g');

%% 高钾玻璃灵敏度分析
x5=[61.71 12.37 5.87 1.11 5.5 2.16 5.09 0.7 1.41];
X5=[x5;x5;x5;x5;x5;x5;x5;x5;x5;x5;x5];
Kfun(x5)
%% 二氧化硅
X51=X5;

```

```

for i=1:11
    X51(i,1)=(0.94+0.01*i)*x5(1);
end
X51=X51';
y51=ones(11,1);
for i=1:11
    y51(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end
%%    氯化钾
X52=X5;
for i=1:11
    X52(i,2)=(0.94+0.01*i)*x5(2);
end
X52=X52';
y52=ones(11,1);
for i=1:11
    y52(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end
%%    氧化钙
X53=X5;
for i=1:11
    X53(i,3)=(0.94+0.01*i)*x5(3);
end
X53=X53';
y53=ones(11,1);
for i=1:11
    y53(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end
%%    氧化镁
X54=X5;
for i=1:11
    X54(i,4)=(0.94+0.01*i)*x5(4);
end
X54=X54';

```

```

y54=ones(11,1);
for i=1:11
    y54(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end

%%    氧化铝
X55=X5;
for i=1:11
    X55(i,5)=(0.94+0.01*i)*x5(5);
end
X55=X55';
y55=ones(11,1);
for i=1:11
    y55(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end

%%    氧化铁
X56=X5;
for i=1:11
    X56(i,6)=(0.94+0.01*i)*x5(6);
end
X56=X56';
y56=ones(11,1);
for i=1:11
    y56(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end

%%    氧化铝
X57=X5;
for i=1:11
    X57(i,5)=(0.94+0.01*i)*x5(7);
end
X57=X57';
y57=ones(11,1);
for i=1:1

```

```

        y57(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end

%%      氧化铝
X58=X5;
for i=1:11
    X58(i,5)=(0.94+0.01*i)*x5(5);
end
X58=X58';
y58=ones(11,1);
for i=1:11
    y58(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end

%%      氧化铝
X59=X5;
for i=1:11
    X59(i,5)=(0.94+0.01*i)*x5(5);
end
X59=X59';
y59=ones(11,1);
for i=1:11
    y59(i)=Kfun(X51(9*(i-1)+1:9*(i-1)+9));
end

%%      整体预测值灵敏度分析
x28=[68.08 2.31 0.26 1.34 1 4.7 0.41 0.33 17.14 4.04 1.04 0.12
0.23 1.96];
X28=[x28;x28;x28;x28;x28;x28;x28;x28;x28;x28];
%      二氧化硅 1
X281=X28;
for i=1:11

```

```

        X281(i,1)=(0.94+0.01*i)*x28(1);
end
X281=X281';
y281=ones(11,1);
for i=1:11
    y281(i)=afun(X281(14*(i-1)+1:14*(i-1)+14));
end
% 氧化钠 2
X282=X28;
for i=1:11
    X282(i,1)=(0.94+0.01*i)*x28(2);
end
X282=X282';
y282=ones(11,1);
for i=1:11
    y282(i)=afun(X282(14*(i-1)+1:14*(i-1)+14));
end

% 氧化钾 (K2O) 3
X283=X28;
for i=1:11
    X283(i,1)=(0.94+0.01*i)*x28(3);
end
X283=X283';
y283=ones(11,1);
for i=1:11
    y283(i)=afun(X283(14*(i-1)+1:14*(i-1)+14));
end

% 氧化钙 (CaO) 4
X284=X28;
for i=1:11

```

```

        X284(i,1)=(0.94+0.01*i)*x28(4);
    end
    X284=X284';
    y284=ones(11,1);
    for i=1:11
        y284(i)=afun(X284(14*(i-1)+1:14*(i-1)+14));
    end
    % 氧化镁(MgO) 5
    X285=X28;
    for i=1:11
        X285(i,1)=(0.94+0.01*i)*x28(5);
    end
    X285=X285';
    y285=ones(11,1);
    for i=1:11
        y285(i)=afun(X285(14*(i-1)+1:14*(i-1)+14));
    end
    % 氧化铝(Al2O3) 6
    X286=X28;
    for i=1:11
        X286(i,1)=(0.94+0.01*i)*x28(6);
    end
    X286=X286';
    y286=ones(11,1);
    for i=1:11
        y286(i)=afun(X286(14*(i-1)+1:14*(i-1)+14));
    end
    % 氧化铁(Fe2O3) 7
    X287=X28;
    for i=1:11
        X287(i,1)=(0.94+0.01*i)*x28(7);
    end
    X287=X287';
    y287=ones(11,1);

```



```

for i=1:11
    y287(i)=afun(X287(14*(i-1)+1:14*(i-1)+14));
end
% 氧化铜 (CuO) 8
X288=X28;
for i=1:11
    X288(i,1)=(0.94+0.01*i)*x28(8);
end
X288=X288';
y288=ones(11,1);
for i=1:11
    y288(i)=afun(X288(14*(i-1)+1:14*(i-1)+14));
end
% 氧化铅 (PbO) 9
X289=X28;
for i=1:11
    X289(i,1)=(0.94+0.01*i)*x28(9);
end
X289=X289';
y289=ones(11,1);
for i=1:11
    y289(i)=afun(X289(14*(i-1)+1:14*(i-1)+14));
end
% 氧化钡 (BaO) 10
X2810=X28;
for i=1:11
    X2810(i,1)=(0.94+0.01*i)*x28(10);
end
X2810=X2810';
y2810=ones(11,1);
for i=1:11
    y2810(i)=afun(X2810(14*(i-1)+1:14*(i-1)+14));
end

```

```

% 五氧化二磷 (P2O5) 11
X2811=X28;
for i=1:11
    X2811(i,1)=(0.94+0.01*i)*x28(11);
end
X2811=X2811';
y2811=ones(11,1);
for i=1:11
    y2811(i)=afun(X2811(14*(i-1)+1:14*(i-1)+14));
end
% 氧化锶 (SrO) 12
X2812=X28;
for i=1:11
    X282(i,1)=(0.94+0.01*i)*x28(12);
end
X2812=X2812';
y2812=ones(11,1);
for i=1:11
    y2812(i)=afun(X2812(14*(i-1)+1:14*(i-1)+14));
end
% 氧化锡 (SnO2) 13
X2813=X28;
for i=1:11
    X2813(i,1)=(0.94+0.01*i)*x28(13);
end
X2813=X2813';
y2813=ones(11,1);
for i=1:11
    y2813(i)=afun(X2813(14*(i-1)+1:14*(i-1)+14));
end
% 二氧化硫 (SO2) 14
X2814=X28;
for i=1:11

```

```

X2814(i,1)=(0.94+0.01*i)*x28(14);
end
X2814=X2814';
y2814=ones(11,1);
for i=1:11
    y2814(i)=afun(X2814(14*(i-1)+1:14*(i-1)+14));
end

plot(0.95:0.01:1.05,y286,'r');
hold on
plot(0.95:0.01:1.05,y282,'b');
hold on
plot(0.95:0.01:1.05,y283,'k');
hold on
plot(0.95:0.01:1.05,y284,'y');
hold on
plot(0.95:0.01:1.05,y285,'g');

```

问题 3

| 类型 | 是否风化 | 二氧化硅 | 氧化钠 | 氧化钾 | 氧化钙 | 氧化镁 | 氧化铝 | 氧化铁 | 氧化铜 | 氧化铅 | 氧化钡 | 五氧化二磷 | 氧化锶 |
|------|------|-------|------|-------|------|------|------|------|------|-------|-------|-------|------|
| 高钾玻璃 | 风化 | 93.96 | 0 | 0.82 | 0.87 | 0.59 | 1.93 | 0.27 | 1.56 | 0 | 0 | 0 | 0 |
| | 无风化 | 67.98 | 2.78 | 10.17 | 6.40 | 1.30 | 6.62 | 2.32 | 2.68 | 0.71 | 1.44 | 1.53 | 0.08 |
| 铅钡玻璃 | 风化 | 38.8 | 2.59 | 0.25 | 2.45 | 1.05 | 4.68 | 0.98 | 0.94 | 34.4 | 9.14 | 5.49 | 0.47 |
| | 无风化 | 24.91 | 1.41 | 0.39 | 2.80 | 1.13 | 2.97 | 0.95 | 2.47 | 43.31 | 13.35 | 5.72 | 0.47 |