# 因子分析法在地球化学测量样品分析中的应用

来源：网络 作者：倾听心灵 更新时间：2023-12-29

*> 摘 要：在地球化学沉积物取样调查过程中，采取的样品的数量极其庞大，潜在控制变量较多，如何快捷有效地提取关键信息，分析数据变化找出异常区域，是样品数据分析的核心问题。利用因子分析方法处理复杂的元素数据，通过计算将庞大的数据划分成少数重要...*

> 摘 要：在地球化学沉积物取样调查过程中，采取的样品的数量极其庞大，潜在控制变量较多，如何快捷有效地提取关键信息，分析数据变化找出异常区域，是样品数据分析的核心问题。利用因子分析方法处理复杂的元素数据，通过计算将庞大的数据划分成少数重要因子，结合区域前期的勘察结果，对因子所代表的古环境进行探讨分析并统计成表、绘制成图，在减少工作量的同时使得数据变得更为直观可靠。为区域地质调查提供了有价值的参考依据，并在后续的勘探过程中得到了验证。

> 关键词：因子分析;地球化学测量;水系沉积物;R型因子

> 1 因子分析的不同方式及其适用范围

因子分析法，其基本目的在于用较少的因子，描述或解释整个事件中变量的关系，不同于主要成分分析，它是通过降维的思想，将原始的研究数据通过矩阵(或协方差矩阵)的形式，以其内部变量关系相互关系为出发点，将错综复杂的变量用少数变量因子来表示的多元统计分析法。

因子分析可根据其出发点不同分为：R型因子分析、Q型因子分析，以及Q-R型因子分析：

R型因子分析，是针对变量所做的因子的分析，其基本思路为通过对变量的相关系数矩阵结构组合的研究，找出能够表现所有变量的少数几个随机变量来描述大多数随机变量之间的相关关系。再根据其相关性的大小对变量进行分组，使同组内的变量之间的相关性较高，而非同组变量之间的相关性较低。

Q型因子分析，是针对样品所做的因子分析。它与R因子的出发点不同，但核心的思路是相同的。它的计算是从样品的相似系数矩阵出发，而R型因子分析的计算是从样品的相关系数矩阵出发的。换而言之就是考虑指标的重要程度，决定保留哪些去掉哪些;Q型聚类分析考虑的是指标之间的相关性，哪几类指标可以合并组成一个类，使得组内距离而小组间距离大(组内距离、组间距离根据具体问题进行具体的调整分析)。

Q-R型因子分析巧妙的运用了R型因子与Q型因子的对偶关系，将变量与样品特性投影在一个因子空间内，使得样品类型特征可以以空间内其他变量点来解释。

水系沉积物样品分析中，核心问题是研究元素共生组合关系，因此R型因子分析被视为样品分析的有效手段和方法，就像原子内部规律组合的结构，R型因子将其所代表的多项变量因子进行整合，实现降维，不仅能体现出元素在含量上的相互关系，更可以反应元素内部成因关系。换而言之，R型因子分析结果可以用来进一步的分析元素的赋存状态和形成方式。在地质事件过程中，经常会伴随物质成分的活化和转移，从而形成新的元素关系网，并在因子分析过程中得到体现。

Q型因子在计算过程中会对指标进行缩减，将保留的重要指标用以代表整体变量，排除分析的过程中因为更大的误差和异常的数值而造成不必要的干扰，并减少了计算量。但在样品分析的过程中，异常值往往是特殊构造、特殊解释的突破口，为了保证实验分析的准确性，往往不采用Q型因子分析法。

Q-R型因子分析的主要应用于研究变量和样品之间的关系，由常用的R型因子分析，可以轻易的转化成Q-R型因子分析，但Q-R型因子分析较少被提及应用不广泛。

因子分析法在样品分析中已经有了实质性的应用，尤其针对土壤、水系地球化学测量过程中分析元素种类庞大，图件数量庞大，单凭人力很难快速准确地进行数据处理分析。相对于其他方式的普查，水系沉积物、土壤沉积物地球化学勘探有着取样简单、成本低、适用范围广等优势。而水系沉积物的成分、含量特征与物源岩性存在着复杂的关系，数据处理是整个地球化学勘探的核心部分，有效的分析方法不仅可以提高样品分析的速度，更可以提高沉积物样品的准确度，对地球化学测量有着重要的意义。

> 2 因子分析法在实际中的应用

在某长江中下游地区的1：20万普查勘探过程中，针对3025个水系沉积物样品进行了28种化学元素的分析。并利用R型因子对元素进行降维，提取有效的公共因子，根据少数公因子提供的变量，来反映不同元素之间的组合关系，进而划分元素共生组合类型。在该地区的水系沉积物样品分析过程中，将28种元素中信息的重叠部分进行组合，提取成为公共因子，以这少数变量综合表现多个变量(此划分基础建立在原样品中具有较多的共同特征)。在实际使用过程中使用了Bartlett球度检验以及KMO检验，在准备过程中对样品进行了相关性检验，对样品相关KMO值进行分析，对实验所得KMO数值进行划分：KMO0.9非常合适;0.8 对区域内28种元素进行分析后，利用正交旋转因子在和矩阵对元素进行整合，将计算结果特征值进行综合对比，最终选取前五个公共因子为重要因子(其特征根数值分别为R1-2.354、R2-1.521、R3-1.358、R4-1.186累计特征根百分比为75%)。这五种重要因子代表工作区内五中元素组合：(1)As，Sb，Mo;(2)Pb，Ag，Bi;(1)ZN，Cu;(4)W;(5)Au。上述五种聚类结合实际数据可以对各种元素之间的亲疏关系进行识别。

聚类分析中所得到的R1组合对应Mo、As、Sb因子组合，代表了高温热液活动的特征。R2组合对应Pb、Ag、Bi因子组合，R3组合对应As、Sb、Mo因子组合，对应为多金属矿化的物质属性，说明上游地区有较为活跃的热液活动。而通过实际勘探已经发现上游地区多为温热型矿床，基本与实验数据相符。R4组合对应KMO计算结果中的W因子，通过两方面证明了W元素具有较强的独立性。R5组合对应Au因子，充分说明Au元素独立的元素特征。

> 3 结束语

利用因子分析所得出的结论，将区域内28种元素减少为5个公共因子，极大的减少了工作量，使得原本无从下手的大数据处理简化成了5个因子的组合，进而转化为较为直观的数据表。将数据加以解释处理按照其分布梯度绘制成图，即可将原数据70%多的信息简单的呈现在了眼前。在实际应用中，图件的辨识度极高，异常明显，元素共生伴生关系便于解释，为下一步的工作打下了良好的基础。

> 参考文献

[1]董毅，范丽琨，段焕春，等.青海大坂山地区水系沉积物测量元素组合分区[J].地质与勘探，202\_，45(1)：70-74.

[2]董毅.因子分析在水系沉积物测量地球化学分区中的应用探讨――以青海都兰地区为例[J].矿产与地质，202\_，22(1)：78-82.

[3]于林松，邱成贵，刘伟.青海省北部绿草山地区化探元素组合分类信息探讨[J].山东国土资源，202\_，30(20)：54-57.

[4]赵博.几种统计分析方法在化探数据处理中的应用[D].北京：中国地质大学，202\_.

本DOCX文档由 www.zciku.com/中词库网 生成，海量范文文档任你选，，为你的工作锦上添花,祝你一臂之力！