

## 地球化学背景值及异常下限确定

确定地球化学背景值与异常下限的方法有很多种。早期采用简单的统计方法求平均值与标准偏差；用直方图法确定的众值或中位数作为地球化学背景值。以后又发展到用概率格纸求背景值与异常下限等。随着对地球化学背景认识的加深，采用求趋势面或求移动平均值等方法来确定背景值和异常下限，70年代以来，多元回归法、稳健多元线性回归分析法、克立格法、马氏距离识别离散点群法等多种方法常作来研究地球化学的背景值和异常下限。

考虑到方法的实用性、有效性、易操作，通过几种方法在工作区的试验对比，迭代法确定的背景值及异常下限较低，更有利于突出弱异常。因此，工作区背景值和异常下限的确定选用迭代法。

迭代法处理的步骤：①计算全区各元素原始数据的均值( $X_1$ )和标准偏差( $Sd_1$ )；②按  $X_1+nSd_1$  的条件剔除一批高值后获得一个新数据集，再计算此数据集的均值( $X_2$ )和标准偏差( $Sd_2$ )；③重复第二步，直至无特高值点存在，求出最终数据集的均值( $X$ )和标准偏差( $Sd$ )，则  $X$  做为背景值  $C_0$ ， $X+nSd$  ( $n$  根据情况选 1.5 或 2, 3) 做为异常下限  $C_a$ 。

采用迭代法求出工作区各地球化学元素特征值及各参数（见表1）。

表1 工作区元素地球化学特征值及参数表

元素	均值 ( $X$ )	标准偏差 ( $Sd$ )	异常下限 ( $X+2Sd$ )	异常下限 ( $X+1.5Sd$ )	$<X+2Sd$ (个数)	$\geq X+2Sd$ (个数)	$\geq X+2Sd$ (%)	迭代 次数
Cu	16.9769	6.83028	30.6375	27.2224	5888	1071	15.39	13
Pb	26.4004	6.25436	38.9091	35.7819	5919	1040	14.94	10
Ag	0.05876	0.016167	0.0910936	0.0830102	5611	1348	19.37	11
Zn	62.0878	17.9671	98.022	89.0384	6049	910	13.08	8
As	4.29365	1.80894	7.91153	7.00706	4803	2156	30.98	19
Sb	0.447151	0.192886	0.832922	0.736479	5767	1192	17.13	11
Hg	0.013682	0.006728	0.0271379	0.023774	6430	529	7.60	6
Au	0.555449	0.277578	1.11061	0.971817	5780	1179	16.94	9

化探数据是以多元素或多变量为特征的。化探数据处理既研究元素之间的相互关系，又研究样品之间的相互关系，前者叫做  $R$  方式分析，后者叫做  $Q$  方式分析。分析结果是将数据按变量或按样品划分成若干类，使各类内部性质相似而各类之间性质相异。如果参加分析的数据含有已知类别（如矿或非矿的作用）能起训练组作用时，数据处理的结果可给出明确的地质解释，否则所做的地质解释就含有较大程度的推测性。

在特定情况下地球化学数据可能只反映单一的地质过程，这样的化探数据是所谓“来自一个母体”的。一般情况是几种地质过程作用在同一地区，他们相互重叠或部分重叠，这反映在地球化学数据上就具有“多个母体”的特征。化探数据处理需要鉴别和分离这些母体，即对化

探数据值进行分解,确定出不同母体的影响在数据中所产生的分量。在确定和分离地球化学母体时常常涉及化学元素的分布形式,如正态分布或对数正态分布等。

地球化学元素的异常下限值确定是地球化学中重要的问题之一,目前还没有一个令人满意的具有科学依据的计算方法。传统的化探异常下限值计算是基于元素的地球化学分布呈正态分布或元素含量在空间上呈连续的变化这一假设为基础的,而事实上地球化学元素含量的空间分布是极其复杂的,研究表明,地球化学景观可能是一个具有低维吸引子的混沌系统,

**相关分析:**相关分析是对客观现象具有的相关关系进行的研究分析。其目的在于帮助我们对关系的密切程度和变化的规律性有一个具体的数量上的认识,作出判断,并且用于推算和预测。其主要内容包括(1)确定现象之间有无关系(2)确定现象之间关系的密切程度(3)测定两个变量之间的一般关系值(4)测定因变量估计值和实际值之间的差异。

**聚类分析:**聚类分析(Cluster analysis)是根据事物本身的特性研究个体分类的方法,其原则是同一类中的个体有较大的相似性,不同类的个体差别比较大。根据分类对象的不同分为样品聚类和变量聚类。

**判别分析:**判别分析是根据表明事物特点的变量值和它们所属的类求出判别函数,根据判别函数对未知所属类别的事物进行分类的一种分析方法,与聚类分析不同,它需要已知一系列反映事物特性的数值变量值及其变量值。

**因子分析:**因子分析是将多个实测变量转换为少数几个不相关的综合指标的多元统计分析方法。在各个领域的科学研究中往往需要对反映事物的多个变量进行大量的观测,收集大量数据以便进行分析寻找规律。多变量大样本无疑会给科学研究提供丰富的信息,但也在一定程度上增加了数据采集的工作量,更重要的是在大多数的情况下,许多变量之间可能存在相关性而增加的分析问题的复杂性,由于个变量之间存在一定的相关关系,因此有可能用比较少的综合指标分别综合存在于各个变量中的各类信息,而综合指标之间彼此不相关,即各指标代表的信息不重叠。这样就可以对综合指标根据专业知识和指标所反映独特含义给予命名,这种方法成为因子分析。

**回归分析:**研究变量之间存在但又不确定的相互关系以及密切程度的分析叫做相关分析,如果把其中的一些因素作为自变量,而另外一些随自变量变化而变化的变量作为因变量,研究他们之间的非确定因果关系,就是回归分析。

**生存分析:**生存分析广泛应用于生物医学,工业,社会科学,商业等领域,例如肿瘤患者经过治疗后生存的时间、电子设备的寿命、罪犯假释的时间、婚姻持续的时间、保险人的索赔等。生存分析就是处理搜集来的数据,生存数据包括生存时间以及相关因素。

**方差分析:**方差分析是检验两个或多个样本均数间差异是否具有统计意义的一种方法,例如:医学界研究几种药物对某种疾病的疗效;农业研究土壤,肥料,日照时间等因素对某种农作物产量的影响;不同饲料对牲畜体重增长的效果等,都可以使用方差分析方法来解决。其基本原理是认为:不同处理组的均数间的差别基本来源于随机误差和实验条件。