

异常下限与异常识别之现状

史长义

(地矿部物化探研究所)

摘要 本文简单介绍了目前用于确定化探异常下限的各种常规方法及其适应性和局限性,并重点介绍了国外勘查地球化学家对这些常规方法的新认识与新观点,以及为了克服这些经典常规统计学方法的不足与弊病而提出的一些新的非常规统计学方法与技术。

一、前言

地球化学异常的解释与评价是建立在正确确定地球化学异常与背景的基础上的。不能确定合适的异常下限,异常的解释推断就可能出现失误。随着勘查地球化学理论的发展以及实践经验的积累,人们越来越认识到这一问题的重要性。在方法上,从直观的或经验式方法到建立在经典统计学基础上的各种单变量与多变量分析方法,确定异常下限的方法多种多样。特别是最近十几年,当人们逐渐认识到那些建立在经典统计学基础上的各种传统方法的局限性时,尤其是稳健统计学在化探数据处理中应用后,又有不少确定异常下限的新方法新技术被提出,使得异常的识别更具客观性与真实性。但是,迄今为止,仍没有一种能为大家所普遍接受的标准化方法。本文就目前某些学者对异常下限的

一些新认识和确定异常下限及异常识别的新的非常规方法作一简单综述。

二、异常下限的概念

有关异常下限的概念,不同的学者有不同的理解。Hawkes 和 Webb (1962)、Rose 等 (1979) 和 Levinsen (1979) 认为异常下限就是背景正常变化的上限。在第六届国际勘查地球化学学术讨论会上,对异常下限所下的定义是:“通过某种方法所获得的一个实在的数字,用它来识别所希望的那些异常样品是指示矿化的”。显然,这个概念是很狭义的,他所指的仅仅是矿异常。考虑到多元素的原因,在第九届国际勘查地球化学学术讨论会上又重新定义异常下限为:能划分出异常和背景的数值就是异常下限 (Garrett 和 Lestinen, 1984)。这就与前苏联地球化学家所提出的“最低异常值”的概念 (见苏联固体矿产化

[13] F.R.Sieyel,……, 1991, Biogeochemical prospecting for goldbearing quartz veins of the Piedmont, Great Falls, Maryland, J.Geochem. Explor. 1991, 41, (3), 275~289.

[14] F.R.Sieyel,……, 1991, Curatella americana L: a biogeochemical sample medium for the Tilaran Montes del Aguacate gold belt, Costa Rica, J.Geochem. Explor. 41, (1-2), 275~289.

[15] M.M.Cole, 1991, Remote sensin, geobotany and biogeochemistry on detection of Thalang a Zn-Pb-Cu deposit Charters Towers, Queensland, Australia, Applied earth science, 100(1) 1~8.

(责任编辑 张锡濂)

探规范)十分类似,与我国通用的异常下限的概念(化探名词解释,1984,物化探研究所编)也一致。Sinclair (1991)则从勘查地球化学的角度出发,认为异常下限是一个利用它能够从背景中划分出异常样品或异常区的数值或特征值。

三、确定异常下限的方法 及一些新认识

确定异常下限的方法在过去的二十几年中有了长足的发展。确定异常下限的基本思路有三点(谢学锦,1979):(1)在地质地理情况简单、面积较小的地区进行区域化探时,可以在全区计算出单一的异常下限;(2)在地质地理情况复杂、面积较大的地区,就要划分一些子区,分别求得几个异常下限;(3)可把异常下限当作一个连续变化着的地球化学面看待,分别根据每一点上的异常下限值确定有无异常。Sinclair (1991)将常用的确定异常下限的方法归纳为三类:经验式方法、基于模式的主观方法和基于模式的客观方法。经验式方法是指那些依靠勘探者经验选择异常下限的方法,包括曾广泛使用的数据评价表及直方图目视评价法。这种方法带有很强的主观意识,它所强调的是绝对含量。因此,在很简单的情况下,这些方法可能是有效的,而在更常见的复杂情况下,用它们所选取的异常下限在几个勘探者之间的重现性则很差。因而,经验式方法不适合于区域性工作。基于模式的主观方法就是将某种正态统计模型或数学模型用于一系列所选取的地球化学参数的方法,例如平均值加二倍标准离差法和趋势分析法。所有的主观模式法选择异常下限的主要困难在于这些方法的模式选择要涉及这样一个事实,即相对于真实情况而言这些模式总是过于简单化。因此,用任何一种主观模式法所确定的异常下限与最佳异常下限

是否具有连续的和可理解的关系并不总是很清楚的。基于模式的客观方法与主观法的唯一不同点在于客观法是依据数据本身的特性而不是勘探者的任意决定来确定异常下限的。但是,这种客观法也含有一些主观因素,它还需要某些假设模式。概率图纸法和间隙统计法即属于这类方法。但是,这两种方法都需要根据假设的相应概率密度函数形式将总数据集分割成单一子数据集。

上述确定异常下限的传统方法中最常用的是算术平均值加两或三倍标准离差法,尤其是在我国的化探工作者中,这种方法用得最为普遍。这种方法最早是由 Hawkes 和 Webb (1962) 针对呈正态或对数正态分布的单一数据总体而提出并首先为 Lepeitier (1969) 所利用,随后则为广大的勘查地球化学家所接受并广泛使用。但是,这一方法却总是被错误地引证和使用,对任何类型的数据和所有的地球化学测量,全都使用平均值加两或三倍标准离差来确定异常下限 (Govett, 1983)。随着实践经验的积累,人们越来越认识到任意使用这种方法的弊端和局限性。因此,对于这种方法最近国外又有学者不断地提出新看法。Sinclair (1974) 曾指出:虽然在某些情况下这种方法可能合适,但是它却忽略了这样一种事实,即精确地把每一数据集顶部 2.5% 的数据作为异常的前提条件 (Priori reason) 是不存在的。而且,在某些情况下,背景总体和异常总体有相当大范围的重合。这种用算术平均值加两或三倍标准离差作异常下限的方法在过去已经引起了许多错误的解释 (Garrett, 1989), 原因是它不能彻底排除异点 (或离群值、野值, 即 outliers) 的干扰。Rose 等 (1979) 也曾指出两或三倍标准离差的选择是任意的, 没有普遍的适用性。但是, 尽管这种方法的局限性很明显, 然而

它却被广泛使用,以致许多勘查地球化学数据的用户都认为异常下限就等于平均值加两倍标准离差 (Garrett, 1989), 原因就是它的用法简单 (Sinclair, 1991)。地球化学数据,特别是微量元素数据的频率分布是正向偏斜的, Dovett 等 (1975) 认为这是由于样品来自混合母体而造成的。此外,这些数据可能还包含有异点和脉冲噪声,而算术平均值对异点的存在很敏感,所以它不是位置的客观估计值 (Chork, 1988)。根据实测化探数据的实际处理结果, Kurzi (1988) 指出,由于经典统计学的基础假设条件很严格,并且要求一定的数据类型,但对化探数据而言,影响所测元素绝对含量和变化的因素 (包括采样的、分析的、环境的和人为污染等) 是多种多样的,一般不能人为控制。所以,这些统计学假设常常不能反映数据的真实特性。极端偏离数据主体的少数测量值或小额量总体可明显地影响估计量,至少部分掩盖数据的固有信息。

由于地球化学数据的频率分布常是正向偏斜的,因此,在使用参数统计时,通常的做法就是把原始数据进行对数转换以使其接近近似正态分布。对于这种对数转换, Govett (1983) 认为相对于低值 (即背景值) 而言,它具有减弱高值 (即异常值) 重要性的作用。因而在一些类型的数据解释中,这是一种自取失败的做法。

对单元素异常下限,有些学者提出了不同的看法,认为单元素异常下限不能客观地反应实际情况。Lindqvist 等 (1987) 指出,用传统的统计学方法确定的单元素异常下限,既不能揭示也不能突出数据集中内含的有价值的复合信息,还会引人歧途,因为在许多情况下异常下限并不是某个重要元素的浓度值,而是样品中几种元素的相互关系。Esbensen 和 Steenfelt (1987) 也指出人们做了极大的努力试图确

定单元素的异常下限,以便用各类统计学方法颇为武断地划分出地球化学异常,尽管这些异常并不总是与所研究的地质问题有密切关系。

为了能更客观地分离背景与异常,尽量排除异点的干扰,最近有不少学者提出了一些新方法,特别是多元素背景与多元素异常成了热门话题。有些人采用新的方法来确定异常下限,如 Esbensen 和 Steenfelt (1987) 的多元双线性投影系统,谢学锦等 (1986) 的 RESMA 系统。而有些学者则干脆避开异常下限的问题,采用一些统计学和数学方法,直接分离出背景总体和异常总体,如 Garrett (1989) 认为在大规模区域调查中,利用传统的异常下限形式,例如规定每一元素取顶部 2% 的数据作为有进一步研究价值的数、将会列出过多的样品来进行详细的研究,而且,在某些环境中,对多元素测量采取单元素异常下限可能会使明显具外来性质的个别部分无法识别出来。因此,他提出一种 GAIT 方法,不用先验地给出一个删除比例,就可以确定元素地球化学异常。Bjorklund 和 Gustavsson (1987) 将点的直径与金属含量相联系,制作元素含量的点源图,能客观地反映元素的变化模式,而无需确定异常下限。

四、异常识别新方法简介

虽然目前用来识别异常下限的各种方法似乎是为大家所接受的,但是从性质方面讲,它们中却很少有几个是可以依赖的 (Probablistic) (Miesch, 1981)。各种用于识别异常下限的方法也应该能够用来估计其所识别出的异常下限占所需异常下限的可能性大小,因此,他提出使用间隙统计法 (gap statistic) 在识别异常下限的同时估计大于它的那些值属于背景总体的可能性大小,即这个异常下限并不是真正异

常下限的可能性。其基本方法为先将地球化学值形成有序数组并按 $\ln(x-\alpha)$ 或 $\ln(\alpha-x)$ 将其转换, 然后标准化, 以使均值为 0, 方差为 1。可从一具单位面的拟合正态曲线上得到期望频率, 某一被检验间隙的中间点可看作异常下限的估计值, 但该点的值要超过相应的临界值。最后, 从相应的概率值就可以看出由这个异常下限分离出的两个地球化学总体的可能性大小。但是, 如果从背景到反映矿化的较高值的变化趋势是连续的话, 该方法就不能识别出异常下限值。从应用情况看, 这种方法比较适合于那些其它方法无法识别的弱异常。通过与概率图的比较, Stanley 和 Sinclair (1989) 认为在异常下限的选择中最好使用概率图解法, 该方法对解释计算结果有多方面的适用性。但是, 用概率图解法确定异常下限时, 人为因素影响太大, 特别是当出现两个以上拐点时。

单变量累积频率曲线确定的地球化学背景, 可揭示出一些单元素的异常含量 (Olesen 和 Armour-Brown, 1984)。为了揭示这些变量的特殊组合特征, 他们提出一种综合性方法, 把 R-型因子剩余、因子得分和单变量筛选方法结合起来, 用于识别异点和确定背景总体, 效果比单一的单变量方法好。其基础思路是建立正交因子模型, 根据剩余的高低删除异点, 再建立新的模型, 再删除, 如此反复, 直到异点可根据地质异常或元素组合解释为止。通过在 South Greenland 的实际应用表明, 这种方法发现的异点中大部分用单一变量的累积频率曲线都没有发现。

使用像梳滤波 (PFF) 和 Kolmogorov-Smirnov 滤波圈定单元素异常 (Chork 和 Cruikshank, 1984; Howarth 等, 1980) 所过滤出的异常既可清晰地指出已知矿化, 又可指示出几个有希望的地段。其中规定异常下限大于或等于两倍标准离差, 这里的

标准离差是通过 PFF 后求出的。

最近, Stanley 和 Sinclair (1988) 讨论了单变量与多变量分析的关系, 指出在通常的多变量分析中一般不考虑单变量分析的结果, 这样进行后续的多变量分析时忽视可辨认出的单变量模式就会使多变量分析的结果及其解释变得模棱两可, 产生不必要的复杂化。因此, 他们认为应该充分利用单变量分析的结果来设计多变量分析的战略。由此他们提出了一种“构造”多变量分析方法。北不列颠哥伦比亚土壤测量的数据分析结果支持了上述观点。

为了有效地解释地球化学数据就要充分考虑数据的多重总体问题 (Riose 等, 1979)。Stanley 和 Sinclair (1987) 提出的在多元素地球化学测量中识别总体和确定异常样品的客观方法——背景特征法 (BCA), 就是为有明显多重总体叠加存在的情况或可能被认为没有异常的情况 (单峰分布) 而设计的。其基本方法是采用概率曲线图来选择每一相关变量的异常下限, 明确地确定一组低于该异常下限的背景样品, 只利用背景样品来确定探途元素的多元背景回归模式。然后, 将这个模式应用于所有可能是异常的样品。根据该模式能成功地预测出元素含量的样品将被划为背景样品, 而与该“函数”呈不同“反应”的那些样品则被认为是异常样品。该方法的基本假设是所有样品都可划分为异常源和背景源两种情况。实践证明, 该方法可为地球化学数据的解释提供补充依据。Lindqvist 等 (1987) 利用 PCA (主成分分析) 法来确定局部多变量异常下限, 分离出背景模型和异常总体, 以建立多变量的地球化学模型。Esbensen 和 Steenfelt (1987) 利用多元双线性投影系统既能确定区域背景模型, 又能将多元素地球化学异常定量化。

Garrett (1989) 认为, 直到目前为

止, 解决异常识别问题通常采用主成分分析和回归分析为基础的方法, 这两种方法(事实上包括所有方法)都存在一个缺陷, 即正寻找的异点可能严重地歪曲了初始分布, 以至在某些情况下, 真正的异常却没有被发现, 而根本是非异常的样品则似乎成了异常。他提出一种多元累积概率图来解决这一问题, 其图形就是 X^2 图, 即 D^2-X^2 值图。每一个多元素地球化学样品在图上用一个值表示。异常值在图上表现为单个点或点群, 可根据剩余数据以间隙与其它数据分离开来。采用迭代法删除异点, 以求得一个均一的背景总体。最后用“背景”的重心和协方差矩阵将所有找到的异点按 D^2 大小顺序显示出来。按照这种思路, 已建立了计算机处理系统, 即 MVT-GAIT 系统。

考虑到常规经典统计学方法都有严格的基本假设条件, 而它们又常常不能描述化探数据的真实特性, Tukey (1977) 引入稳健统计学提出了一种新的非常规数据处理方法——勘查数据分析 (Exploratory Data Analysis) 技术, 最近几年引起了人们的重视 (Campbell, 1982; Smith et al., 1982; Howarth, 1983, 1984; Kurzl, 1988, 1989; Chork, 1988; 史长义, 1991, 1993)。这种方法不要求任何假设条件, 而是根据数据本身所固有的特性来识别异点, 确定背景总体和异常总体, 且不需要预先对原始数据作任何数学处理 (史长义, 1991, 1993)。EDA 方法中包含一些简单而有效的图示技术, boxplot

(框图) 和 steam-and-leaf (茎-叶) 图就是其中的两个最有用的图示, 采用五参数综合法描述单元素的分布特征比各种单一方法更客观、直观和准确, 同时不受假设条件的影响, 而且框图对比还可以强化解释推断能力。Kurzl (1989) 将 EDA 方法编入计算机资料处理系统, 处理奥地利的区域水系沉积物资料。目前, EDA 方法已为越来越多的用户所使用。

五、结束语

虽然确定异常下限的方法多种多样, 但是, 它们都有各自的适用性和局限性, 尤其是那些建立在常规经典统计学基础上的方法之局限性更大, 其主要原因是化探数据的复杂性使得它们很难满足那些统计学方法的基本假设。这样, 采用这些方法来处理化探数据, 就有可能得出令人无法解释的结果。因此, 为了解决或者避开这一问题, 不少学者提出了一些新的非常规统计学方法。本文旨在介绍国外学者对这一问题的某些新认识、新观点, 以及最近提出的一些新方法、新技术, 供大家在实际工作中参考。特别介绍一些国外学者对国际上, 尤其是国内化探工作者在确定异常下限的方法中使用最多的以“平均值加两或三倍标准离差作为异常下限”的方法的一些新认识, 希望能够引起国内同行的注意。

承蒙谢学锦先生审阅全文并提出宝贵意见, 特此致谢。

参考文献

- [1] Bjorklund A. and Gustavsson N., 1987. Visualization of geochemical data on maps: new options, J. Geochem. Explor., 29: 89-103.
- [2] Campbell N.A., 1982. Statistical treatment of geochemical data. In: R.E. Smith (editor), Geochemical exploration in deep weathered terrain, P141-144.
- [3] Chork C.Y. and Cruikshank B.I., 1984. Statistical map analysis of regional stream sediment data

from Australia. *J. Geochem. Explor.*, 21, 405–419.

[4] Chork C.Y., 1988, Spatial filtering of exploration geochemical data using EDA and robust statistics. 3rd CEGS papers, P263–267.

[5] Esbensen K.H. and Steenfelt A., 1989, Geochemical prospecting in complex sample media — multivariate data analysis of indirect observation. *J. Geochem. Explor.*, 32: 345–348.

[6] Garrett R.G., 1989, The chi-square plot: a tool for multivariate outlier recognition. *J. Geochem. Explor.*, 32:319–343.

[7] Garrett R.G. and Lestinen P., 1984, Thresholds and anomaly interpretation. *J. Geochem. Explor.*, 21: 137–142.

[8] Govett G.J.S., 1983, Rock geochemistry in mineral exploration, Vol.3 of Handbook of Exploration Geochemistry, Elsevier, 462pp.

[9] Govett G.J.S., Goodfellow W.D., Chapman R.P. and Chork C.Y., 1975, Exploration geochemistry — distribution of elements and recognition of anomalies. *Math. Geol.*, 7: 415–446

[10] Hawkes H.E. and Webb J.S., 1962, Geochemistry in Mineral Exploration. Harper and Row, New York, N.Y., 415pp.

[11] Howarth R.J., 1983, Mapping. In: Handbook of geochemical exploration, Vol.2. P111–205.

[12] Howarth R.J., 1984, Statistical applications in geochemical prospecting: a survey of recent developments. *J. Geochem. Explor.*, 21: 41–61.

[13] Howarth R.J., Koch G.S., Chork C.Y., Carpenter R.H. and Schuenemeyer J.H., 1980. Statistical map analysis techniques applied to regional distributions of uranium in stream sediment samples from the Southeastern United States for the National Uranium Resource Evaluation Program. *Math. Geol.* 12: 339–366.

[14] Kurzl H., 1988, Exploratory data analysis: recent advances for the interpretation of geochemical data. *J. Geochem. Explor.*, 30:309–322.

[15] Kurzl H., 1989, Data analysis and geochemical mapping for the regional stream sediment survey of Austria. *J. Geochem. Explor.*, 32: 349–352.

[16] Lepeltier C., 1969, A simplified statistical treatment of geochemical data by graphical representation. *Econ. Geol.*, 64: 538–550.

[17] Levinson A.A., 1979, Introduction to Exploration Geochemistry(Second Edition). Applied publishing Ltd., Wilmette, Illinois, 924pp.

[18] Lindqvist L. Lundholm, I., Niska D., Esbensen K. and Wold S., 1987, Multivariate geochemical modelling and integration with petrophysical data. *J. Geochem. Explor.*, 29:279–294.

[19] Miesch A.T., 1981, Estimation of the geochemical threshold and its statistical significance. *J. Geochem. Explor.*, 16:49–76.

[20] Olesen B.L. and Armour-Brown A., 1984, Multivariate screening of training sets for classification and the definition of geochemical background. *J. Geochem. Explor.*, 21:471–479.

[21] Rose A.W., Hawkes H.G. and Webb J.S., 1979, Geochemistry in mineral exploration (2nd edition), Academic Press, London, 657pp.

[22] Sinclair A.J., 1974, Selection of threshold values in geochemical data using probability graphs.

J.Geochem. Explor., 3:129-149.

[23] Sinclair A.J., 1991, A fundamental approach to threshold estimation in exploration geochemistry: probability plots revisited. J.Geochem. Explor., 41:4-22.

[24] Smith R.E., Campbell N.A., and Perdix J.A., 1982, Identification of some Western Australian Cu-Zn and Pb-Zn gossans by multi-element geochemistry. In: R.E. Smith(Editor), Geochemical exploration in deep weathered terrain. P75-90.

[25] Stanley C.R. and Sinclair, A.J. 1987, Anomaly recognition for multi-element geochemical — A background characterization approach, J.Geochem. Explor., 29:333-353.

[26] Stanley C.R. and Sinclair A.J. 1988, Univariate patterns in the design of multivariate analysis techniques for geochemical data evaluation. P113-130, in Quantitative Analysis of Mineral and Energy Resources, edited by C.F.Chung, A.G.Fabbi and R.Sinding-Larsen, D.Reedel, Dordrecht. 738pp.

[27] Stanley C.R. and Sinclair, A.J. 1989. Comparison of probability plots and gap statistic in exploration of thresholds for exploration geochemistry data. J.Geochem. Explor., 32: 355.

[28] 史长义, 1991, 化探数据解释推断的新方法——EDA 技术, 国外地质勘探技术, 第 1 期, P38-41.

[29] 史长义, 1993, 勘探数据分析 (EDA) 技术的应用, 地质与勘探, 第 11 期, P52-58.

[30] 谢学锦, 1979, 区域化探, 地质出版社, 192pp.

[31] 谢学锦等, 1986, 快速评价与筛选多元异常 (RESMA) 的计算机系统.

(责任编辑: 张锡谦)

用建筑垃圾加固软土地基

大港油田地处渤海湾两岸滨海平原上, 地基土层为滨海相沉积的软弱土层。油田矿区和基地的工业、民用建筑 3~7 层楼房, 不是采用筏片基础就是采用桩基, 致使油田基本建设投资过高。与此同时, 油田内部却有着堆积如山的建设垃圾。这些垃圾不易处理, 占用了大量场地, 还对环境造成严重污染形成公害。为节省基建投资, 变废为宝, 油田在建设若干 4~6 层民用及工业用楼房时, 利用建筑垃圾加固软土地基, 取得了较好的社会、经济效益。

加固软基施工工艺: 用吊车将柱形状重锤提升到一定高度使之自由下落夯击地基。经数击后, 便冲夯成一个直径略大于锤底直径, 深度达 1.5~2.5m 的浅坑孔。用铲车向孔内填充适量的、稍加粉碎的砖和混凝土小块 (粒径 4~6cm) 建筑垃圾填料, 使之打入土中。锤夯次数以使孔内填料能托住重锤为宜。然后再次填料夯击至填满浅孔为止。所形成的坚实的短柱体, 称之为夯填加固柱体。加固柱体在场区内可按矩形、等腰三角形或梅花形布置。按一定顺序完成各加固体后, 场区地基便得到加固, 形成由若干加固体组成的复合地基。为使上部荷载能较均匀地传递到处理后的地基上, 以发挥加固柱体与周围土体的共同作用, 对地基用低落锤普遍地进行夯击, 同时做 30~40cm 厚的三土灰土或建筑垃圾垫层。

加固机理: 由于锤夯的冲击作用使填料向夯击方向及侧向挤密, 从而对其周围的土体产生挤密加固作用, 形成一个自内向外的挤密圈。在挤密过程中, 周围土体的孔隙水压力随之增高, 形成超静孔隙水压力。粘性土的固结所需要的时间和排水距离的平方成正比。排水距离越远, 固结延续的时间越长。加速软土固结的有效方法是增加排水途径, 缩短排水距离。采用夯填建筑垃圾的方法, 可使地基土的超静孔隙水压力通过加固柱体迅速消散, 加快软土固结, 提高复合地基的强度, 满足楼房的承载要求。

(杨和季)

— 25 —