

doi: 10.11720/wtyht.2020.1014

耿国帅,杨帆,郭建娜.ILR 变换后数据的因子分区标准化在东昆仑东段地球化学异常圈定中的应用[J].物探与化探,2020,44(1):112-121.
http://doi.org/10.11720/wtyht.2020.1014

Geng G S, Yang F, Guo J N. The application of ILR transomed data factor analysis to delineating geochemical anomalies[J]. Geophysical and Geochemical
Exploration, 2020, 44(1): 112-121. http://doi.org/10.11720/wtyht.2020.1014

ILR 变换后数据的因子分区标准化 在东昆仑东段地球化学异常圈定中的应用

耿国帅^{1,2}, 杨帆^{3,4}, 郭建娜⁵

(1. 中国地质大学(北京) 地球科学与资源学院, 北京 100083; 2. 中国人民武装警察部队 黄金地质研究所, 河北 廊坊 065000; 3. 北京矿产地质研究院, 北京 100012; 4. 中国地质调查局 土地质量地球化学调查评价研究中心, 河北 廊坊 065000; 5. 河北省廊坊市自然资源和规划局, 河北 廊坊 065000)

摘要: 从区域地球化学勘查数据中提取找矿地球化学信息是目前化探工作的重要研究课题之一。地球化学背景与异常划分是提取找矿地球化学信息的关键, 而地球化学数据是成分数据, 具有闭合效应, 使得数据在进行主成分分析和因子分析时需进行变换。笔者采用东昆仑东段水系沉积物中 11 种常量元素的含量数据, 通过 ILR 数据变换后进行因子分析, 根据因子分区结果, 结合 EDA 法, 对样品数据进行分区标准化, 以标准化后的值来圈定异常。结果表明, 与传统方法相比, 利用 ILR 变换后的因子分区标准化方法所圈定的异常可以去除地质体中高背景值的影响, 与矿床空间位置对应较好, 且该方法受人为因素影响较小, 实际应用效果显著。

关键词: 区域化探; ILR 变换; 因子分析; EDA 法; 异常圈定; 东昆仑

中图分类号: P632 **文献标识码:** A **文章编号:** 1000-8918(2020)01-0112-10

0 引言

地球化学背景与异常划分是地球化学找矿中信息提取的关键问题。元素地球化学背景受地质环境、景观条件等诸多因素的制约, 在一定区域内、同一景观条件下, 地质环境是影响元素地球化学背景的主要因素^[1-3]。以往按 1:20 万图幅确定统一异常下限的做法, 导致低背景区矿化信息被掩盖, 而高背景区出现了较多的非矿异常, 背景与异常的划分存在一定的不合理性。为此, 勘查地球化学工作者使用移动平均法、分区背景校正法^[4]、子区中位数衬值滤波法^[5-6]、归一化法^[7]、小波分析法^[8]、趋势

面法^[9-10]、C 型转换法^[11]、泛克里格法和局部奇异性分析法^[12]等来处理区域化探数据, 在很大程度上提高了异常找矿信息的可靠性, 有其优越性, 但仍然存在一定的局限性。这些方法要么注重控制元素含量分布的地质因素, 要么把地球化学场看作一个连续变化的曲面而采用各种方法去拟合。

时艳香提出了地球化学单元的概念, 利用 R 型因子分析方法对样品进行地球化学分区^[13], 并在实际应用中取得了一定的效果, 使地球化学分区标准化得到了很好的应用和推广^[14-19]。因子分区标准化的方法充分考虑了地质背景与元素组合之间的关系, 有一定的客观性。但地球化学元素的含量数据是成分数据, 具有闭合效应。Filzmoser 指出, 地球化学数据在进行计算时, 如果不对数据进行变换, 即使

收稿日期: 2019-01-07; **修回日期:** 2019-07-25

基金项目: 国家自然科学基金项目(41773030;41373048); 中央级公益性科研院所基本科研业务费专项项目(AS2015J03)

作者简介: 耿国帅(1972-), 男, 高级工程师, 在读博士研究生, 主要从事勘查地球化学相关基础理论研究和调查评价工作。Email: hnsxmgs@163.com

通讯作者: 杨帆(1983-), 男, 高级工程师, 博士研究生, 主要从事勘查地球化学及环境地球化学相关基础理论研究和调查评价工作。Email: yf1983@163.com

是做平均值和标准方差这样简单的统计或者是做直方图,都可能得出错误的结果;用原始数据统计标准差和方差几乎是没有意义的^[20],方差矩阵是做因子分析和主成分分析的基础,方差矩阵的不准确会使因子分析的结果与实际结果不相符,从而影响其效果。因此,笔者试图采用 ILR 变换后的因子分区标准化法来圈定地球化学异常,旨在进一步消除高背景地质体对异常圈定的影响。结果表明,与传统方法相比,利用 ILR 变换后的因子分区标准化方法所圈定的地球化学异常可以去除地质体中高背景值的

影响,与矿床空间位置对应较好,且该方法受人为因素影响较小,实际应用效果显著。

1 研究区概况

青海省东昆仑东段跨越了古亚洲和特提斯两大构造域,位于南北两大构造域结合部位,区内可划分为秦祁昆、特提斯 2 个一级单元,东昆仑造山带、巴颜喀拉造山带 2 个二级构造单元及 4 个三级构造单元(表 1,图 1)。

表 1 东昆仑东段大地构造分区		
Table 1 Division of geotectonic classification in the east of Dongkunlun		
一级构造单元	二级构造单元	三级构造单元
秦祁昆(东昆仑—祁连—北秦岭)晚加里东造山系(I)	I ₂ 东昆仑造山带	I ₂ ⁶ 祁漫塔格—都兰造山亚带(昆北带); I ₂ ⁷ 伯喀里克—香日德元古宙古陆块体(昆中带); I ₂ ⁸ 雪山峰—布尔汉布达造山亚带(昆南带)
特提斯(东特提斯北部)华力西—印支造山系(II)	II ₁ 巴颜喀拉晚印支造山带	II ₁ ¹ 布喀达坂峰—阿尼玛卿华力西、印支复合造山亚带(北巴带)

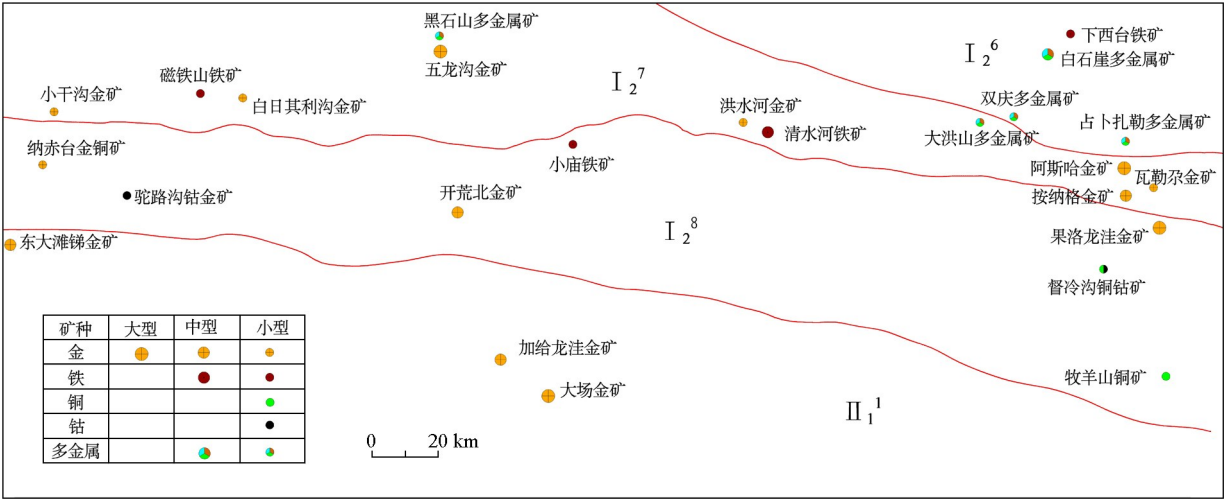


图 1 东昆仑东段矿床分布及构造分区

Fig.1 The sketch map of deposit and geotectonic in study area

不同的沉积环境具有不同的岩石组合,研究区昆北主要为海相沉积的碳酸盐岩和火山岩组合;昆中主要为古老变质岩和花岗岩组合,昆南主要为洋盆环境下形成的火山岩、碎屑岩和碳酸盐岩组合,而北巴则主要为三叠纪的复理石沉积建造。不同的岩石组合造成了不同的地球化学特征和矿床组合。

2 成分数据的概念和变换

成分数据是一组部分占整体的比例数据,只携带相对信息。目前成分数据不再被认为是具有定和的比值向量,而扩展为比值向量的等价类。一个成

分数据是一个代表任意元素组合的等价类,其分量的和可以不是一个常数。元素含量作为相对含量,其和是一个常数,是典型的成分数据,各元素间不是独立的,而是相互影响的,这种影响对常量元素尤其严重,例如样品中 SiO₂ 含量的增加,会造成其他元素含量的总体降低,这就会造成 SiO₂ 与其他元素间通常呈负相关关系。因此在数据处理之前,如果不进行一定的变换,就可能得出错误的结论,尤其是相关性计算和多元统计分析^[21-24]。

目前成分数据的处理与分析有 3 种变换:ALR 变换(additive log-ratio transformation)、CLR 变换(centered log-ratio transformation)和 ILR 变换(iso-

metric log-ratio transformation)。3 种变换的共同点在于变换形式都是变量的对数比值,因而统称为对数比变换。本质上,ALR 变换和 CLR 变换不能消除成分数据的闭合效应,而 ILR 变换可以;原始成分数据经 ALR 变换和 ILR 变换之后,减少一个变量;CLR 变换之后的数据与原始数据的变量数相同。

从 3 种变换的结果来看,ILR 变换无疑是最好的,但由于变换结果减少了一个变量,增加了结果解释难度,而 CLR 变换能够保持变换前后变量数不变,因此 Filzmoser 利用标准正交基,把 ILR 变换后的因子分析结果在 CLR 空间下表示出来,既克服了成分数据的闭合效应,又突显了元素组合的地球化学意义。CLR 变换和 ILR 变换以及与之联系的标准正交基简要介绍如下。

以一个原始数据矩阵 X_{ij} ($m \times n$, 其中 m 为样品数, n 为元素个数) 为例,分别介绍 2 种变换。

1) CLR 变换

$$\text{clr}(\mathbf{X}) = \ln \left[\frac{x_{1j}}{g(x_1)}, \frac{x_{2j}}{g(x_2)}, \dots, \frac{x_{mj}}{g(x_m)} \right],$$

$j = 1, 2, \dots, n; g(x_i)$ 为第 i 个样品的几何平均值。

2) ILR 变换

ILR 变换是在 CLR 变换和标准正交基 v_j 的基础上得来的,标准正交基^[25]

$$v_j = \sqrt{\frac{j}{j+1}} \left(\frac{1}{j}, \dots, \frac{1}{j}, -1, 0, \dots, 0 \right)';$$

$$j = 1, 2, \dots, n - 1。$$

$$\text{ilr}(\mathbf{X}_j) = \sqrt{\frac{j}{j+1}} \ln \frac{\sqrt{\prod_{k=1}^j x_k}}{x_{j+1}};$$

$$j = 1, 2, \dots, n - 1。$$

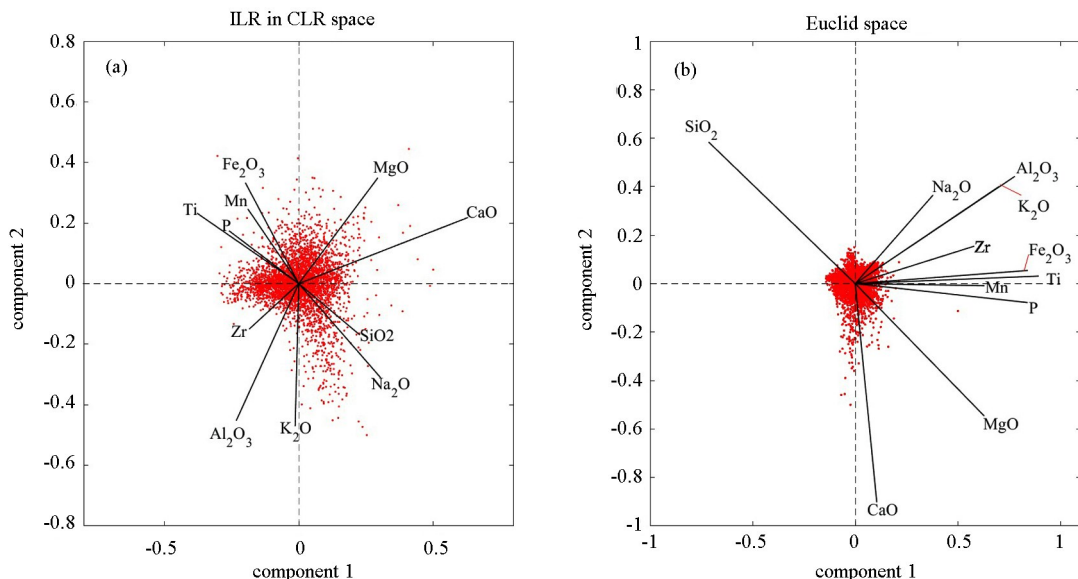
因此 CLR 变换和 ILR 变换可以通过标准正交基联系起来。

$$\text{clr}(\mathbf{X}) = \mathbf{V} \times \text{ilr}(\mathbf{X}), \text{ilr}(\mathbf{X}) = \text{clr}(\mathbf{X}) \times \mathbf{V},$$

$$\mathbf{V} = [v_1, v_2, \dots, v_{n-1}]。$$

笔者利用研究区所收集的 4 001 件水系沉积物样品中 11 种常量元素(或氧化物)的含量数据,将其未经变换和经 ILR 变换后的数据作主成分分析(图 2)。结果表明,未经 ILR 变换的元素(或氧化物)主要分布在主成分双标图的右半部分,除 SiO_2 , 其余 10 个元素或氧化物在第一主成分的载荷都是正数;在第二主成分中,除 CaO 、 MgO 、 P 、 Mn 等元素和氧化物的载荷为负外,其余元素和氧化物的载荷为正; SiO_2 和其他元素间隔较远,显示 SiO_2 与其他元素的相关性较差,成独立一簇。

对数据作 ILR 变换,计算其主成分并变换在 CLR 空间下显示其结果。经 ILR 变换的主成分双标图显示,11 个元素(或氧化物)较均匀地分布在一~四号象限中,第一象限中是 CaO 和 MgO ,代表钙碱性元素的氧化物,第二象限 Fe_2O_3 、 Mn 、 P 、 Ti 等,代表的是基性元素组合,第三象限是 Al_2O_3 、 K_2O 氧



a—经 ILR 变换后的第一、第二主成分双标图;b—未经变换的第一、第二主成分双标图;图中红色圆点为第一和第二主成分的得分
a—the biplot of the first and second principal component from ILR transformed geochemical data;b—the biplot of the first and second principal component from raw geochemical data;the red dots in the picture are the scores of the first and second principal components

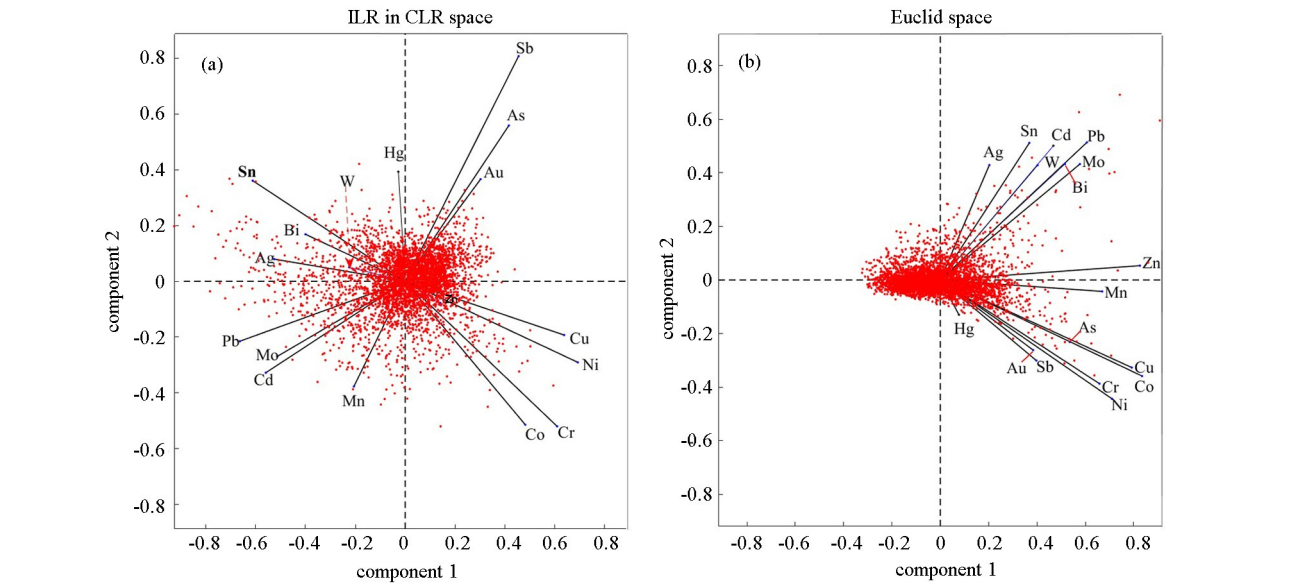
图 2 青海省东昆仑东段常量元素 ILR 变换和未经变换的主成分分析双标图

化物,代表碱性氧化物,而第四象限是 SiO₂、Na₂O,代表的是酸性元素组合。

同时,笔者又利用 Ag、As、Au、Bi、Cd、Co、Cr、Cu、Hg、Mo、Ni、Pb、Sb、Sn、W、Zn 等 17 种成矿元素经 ILR 变换和未经变换的数据分别做主成分分析双标图,结果显示,未经变换的数据双标图中,17 种元素都分布在第一主成分的正载荷部分,而经变换后

的数据,则较均匀地分布在各个象限(图 3)。

综上对比结果可知,对原始化探数据作主成分分析有 2 个重大缺陷:①由于成分数据存在闭合效应,原始化探数据的第一主成分的载荷几乎都为正数,且这一现象普遍存在;②承载最大信息量的第一主成分通常是难以解释的,而经过 ILR 变换后的第一主成分则具有较为明确的地球化学意义。



a—经 ILR 变换后的第一、第二主成分双标图;b—未经变换的第一、第二主成分双标图;图中红色圆点为第一和第二主成分的得分
a—the biplot of the first and second principal component from ILR transformed geochemical data;b—the biplot of the first and second principal component from raw geochemical data;the red dots in the picture are the scores of the first and second principal components

图 3 青海省东昆仑东段成矿元素经 ILR 变换和未经变换的主成分分析双标图

Fig.3 Biplot of principle component analysis from the ore forming elements in the study area

3 ILR 变换的数据因子分区标准化

主量组分是构成岩石的主体,主量组分的组成和相对含量决定了岩石性质,支配了微量元素的地球化学行为及其在岩石中的分配,因此主量组分可

3.1 数据处理步骤

首先,选择研究区水系沉积物样品中常量元素的氧化物(Al₂O₃、CaO、Fe₂O₃、K₂O、MgO、Na₂O、SiO₂)和含量相对较大的元素(Mn、P、Ti、Zr)共 11 种成分,并对数据进行 ILR 变换。其次,确定分区类型,即对样品中所选择的成分作因子分析,根据因子载荷矩阵中所反映的不同组合来确定分区的地球化学类型。然后,确定子区边界,因子得分是标准化值,可以进行大小比较。当某样品在因子 k 中的得分值的绝对值最大,说明该样品在元素组合 k 中所

占份额最大,故将其归为第 k 类(其中正值为 k_1 类,负值为 k_2 类)

根据分区结果对各变量进行分区标准化,圈定地球化学异常。标准化公式如下:

$$Z_{ik} = \frac{X_{ik} - Q_{3k}}{1.5 \times IQR_k}。$$

其中 Z_{ik} 表示标准化后的值; X_{ik} 为原始分析值; Q_{3k} 为第 k 类样品的上四分位值; IQR_k 为第 k 类样品的内散度,其值为上分四位点的值减去下四分位点的值。

分区标准化是基于 EDA(exploratory data analysis)非常规化探数据处理方法,EDA 技术以稳健统计学为基础,不需要设定假设条件,也无需对原始数据作任何处理,而是根据数据本身的特点来提取异常。EDA 方法的统计参数包括中位数、2 个极值点和上下四分位点,该方法认为数据集中有 25% 的数据是无效的,这些数据对整个数据集的中位数和上下四分位点不产生影响。绝大多数情况下,EDA 方

法提供的中位数比平均值能更好地估计数据的中心位置,内散度比标准离差能更好地估计数据的离散程度。它们都能抵抗特高值和特低值的干扰,比传统的方法能更有效地处理化探数据,因此 EDA 法受到了地球化学家的关注^[26-32]。

根据 EDA 方法以 $Q_3+1.5$ 倍 IQR 作为地球化学异常下限,并根据上述标准化计算公式,就能很方便地圈定异常,即标准化值大于 1、小于 2 是弱异常点,大于 2 是强异常点。

3.2 结果分析

基于主成分分析方法进行(初始)因子提取,首先对 ILR 变换后的 11 种成分数据进行主成分分析,计算出相关系数矩阵的特征根和解释率(见表 2);根据因子特征值(>1)和累计解释率(>75%)的要求,选择前 4 个主成分作为主因子,进行因子旋转(见表 3);然后计算各样品的主因子得分,制作各主因子得分计量图(图 4);根据各样品因子得分绝对值的大小,确定样品的地球化学类型,制作地球化学分区图(图 5)。

由表 3、图 4 和图 5 可知, $F1_1$ 为 P、Ti、Zr、 Fe_2O_3 组合,反映的是与基性、超基性岩分布有关的元素组

合,主要分布于昆南带内,此外还分布于昆南断裂带上,与该区中基性岩的分布较一致; $F1_2$ 为 Na_2O 、CaO、 SiO_2 、 K_2O 组合,反映的是与中酸性岩有关的元素组合,主要分布于昆北及昆中带上,与中酸性花岗岩的分布范围较一致。 $F2_1$ 为 CaO、MgO 组合,反映的是与生物沉积为主的灰岩、白云岩等元素组合,分布于昆中和昆南带上,与碳酸盐岩的分布范围较一致; $F2_2$ 为 Al_2O_3 、 K_2O 组合,反应的是与碎屑沉积为主的砂泥岩的元素组合,分布于北巴颜喀拉,与北巴颜喀拉地层的板岩分布范围较一致。 $F3_1$ 为 K_2O 、MgO、 Na_2O 组合,反映的是与河流的蒸发作用有关的元素组合,主体分布于柴达木盆地南缘; $F3_2$ 为 SiO_2 、Zr 组合,反映的是与河流的搬运作用有关的元素组合,主体分布于昆南断裂带以南。 $F3_1$ 和 $F3_2$ 是表生地球化学作用因子,在柴达木盆地的南缘,地表由于蒸发作用,可溶性组分结晶形成各种盐类,形成 K_2O 、MgO、 Na_2O 等易溶元素组合,东昆仑断裂带以南,蒸发作用弱,而与河流的搬运为主,形成难溶的 Si、Zr 等元素组合。 $F4_1$ 为 Na_2O 、 K_2O 、Zr、CaO 组合,反映的是与河流碎屑沉积为主的元素组合,主要分布于昆中带上,与河流碎屑沉积环境有关; $F4_2$

表 2 研究区主成分分析特征值和累积方差贡献率

Table 2 The eigenvalue and variance explained of principal component analysis in the study area			
主成分	特征值	方差贡献率/%	累计贡献率/%
PC1	2.95429	29.5429	29.5429
PC2	2.209107	22.09107	51.63397
PC3	1.451425	14.51425	66.14822
PC4	1.26463	12.6463	78.79452
PC5	0.686685	6.866845	85.66137
PC6	0.523345	5.233451	90.89482
PC7	0.461487	4.614866	95.50969
PC8	0.252381	2.523811	98.0335
PC9	0.11301	1.130104	99.1636
PC10	0.08364	0.8364	100

表 3 研究区正交旋转因子载荷

Table 3 Orthometric rotating factor loading matrix of factor analysis in the study area				
指标	F1	F2	F3	F4
Mn	-0.07738	-0.07873	-0.09688	-0.83041
P	0.702332	-0.01513	0.075709	0.038553
Ti	0.710837	-0.14553	0.096094	-0.26549
Zr	0.482694	-0.18209	-0.38805	0.396661
Al_2O_3	-0.2006	-0.8244	-0.00632	0.018327
CaO	-0.50636	0.941437	0.227767	0.310812
Fe_2O_3	0.215612	0.058648	0.045556	-0.63857
K_2O	-0.29089	-0.59256	0.308376	0.3833
MgO	-0.08812	0.649321	0.456313	-0.03996
Na_2O	-0.5376	-0.03472	0.228541	0.462876
SiO_2	-0.41052	0.223752	-0.9471	0.163893

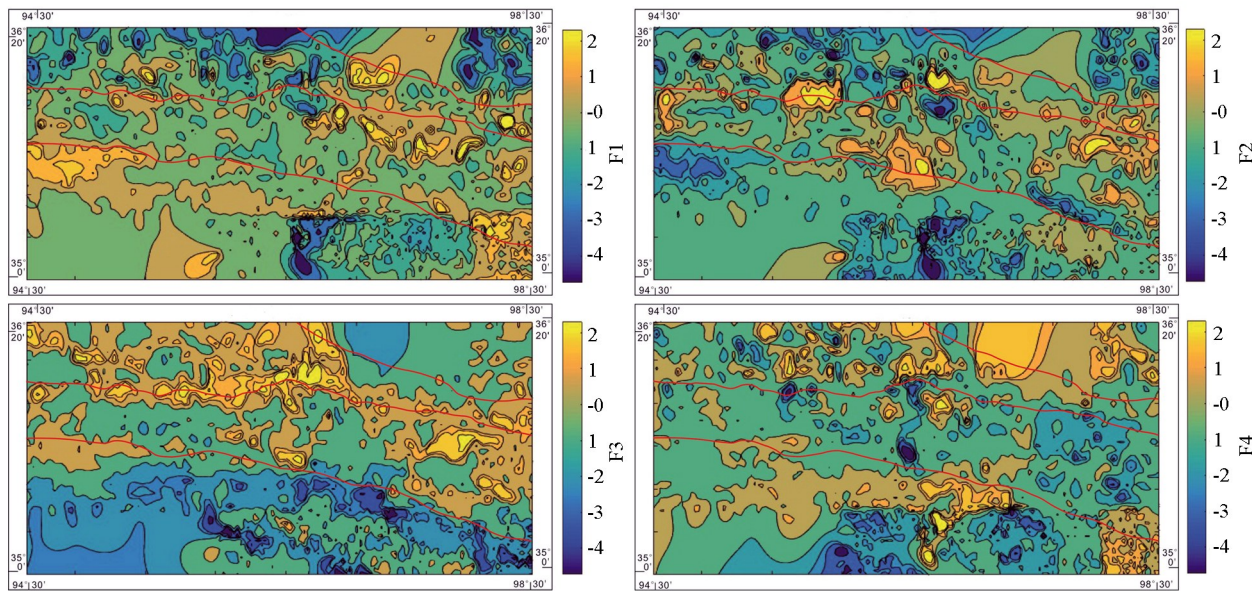


图 4 研究区常量元素经 ILR 变换后 4 个因子得分计量

Fig.4 Four factor score maps of major component from ILR-transformed data in study area

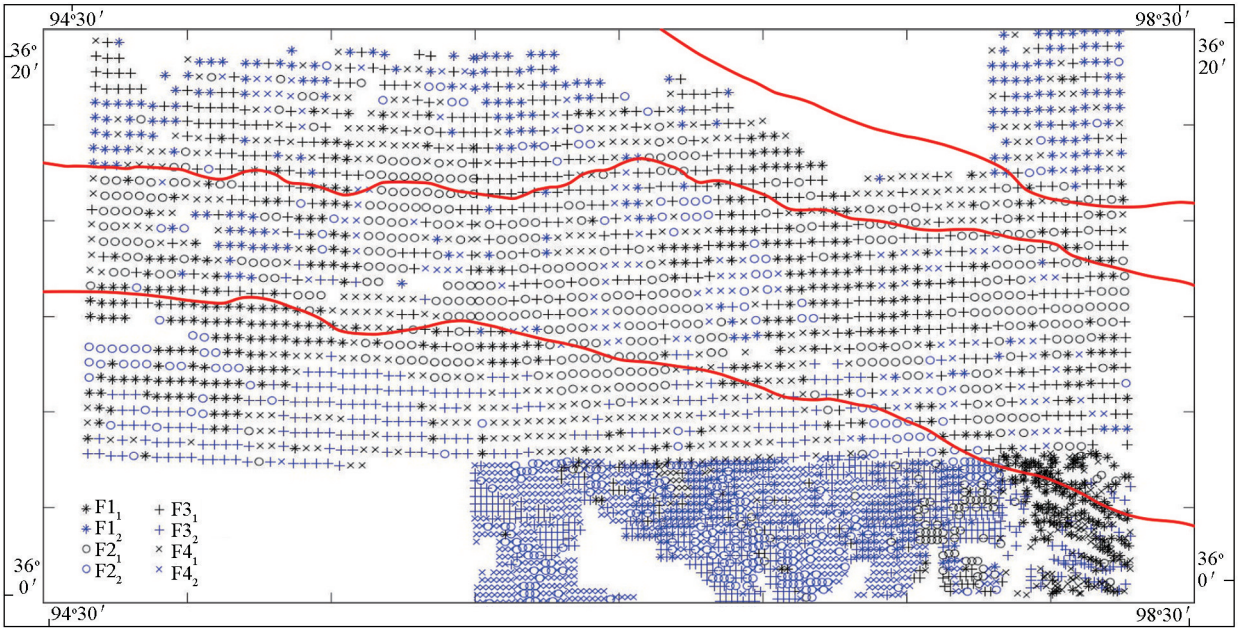


图 5 研究区地球化学分区

Fig.5 Geochemical subdivision map in the study area

为 F_2O_3 、Mn 组合,反映的是与海相页岩、泥质岩为主的元素组合,主要分布于北巴,与海相的泥、页岩沉积环境有关。

为了解各地球化学分区中常量元素的含量分布,统计了 8 类地球化学分区中各常量组分的中位数(表 4)。结果显示,P、Ti、Zr 中位数的最大值都处于 F_{1_1} 中,而最小值都处于 F_{1_2} 中;CaO 中位数的最大值处于 F_{2_1} 中,而最小值处于 F_{2_2} 中; Al_2O_3 、 K_2O 中位数的最大值都处于 F_{2_2} 中; Na_2O 、MgO 中位数的最大值及 Al_2O_3 、 K_2O 的次大值都处于 F_{3_1} 中,而最小值

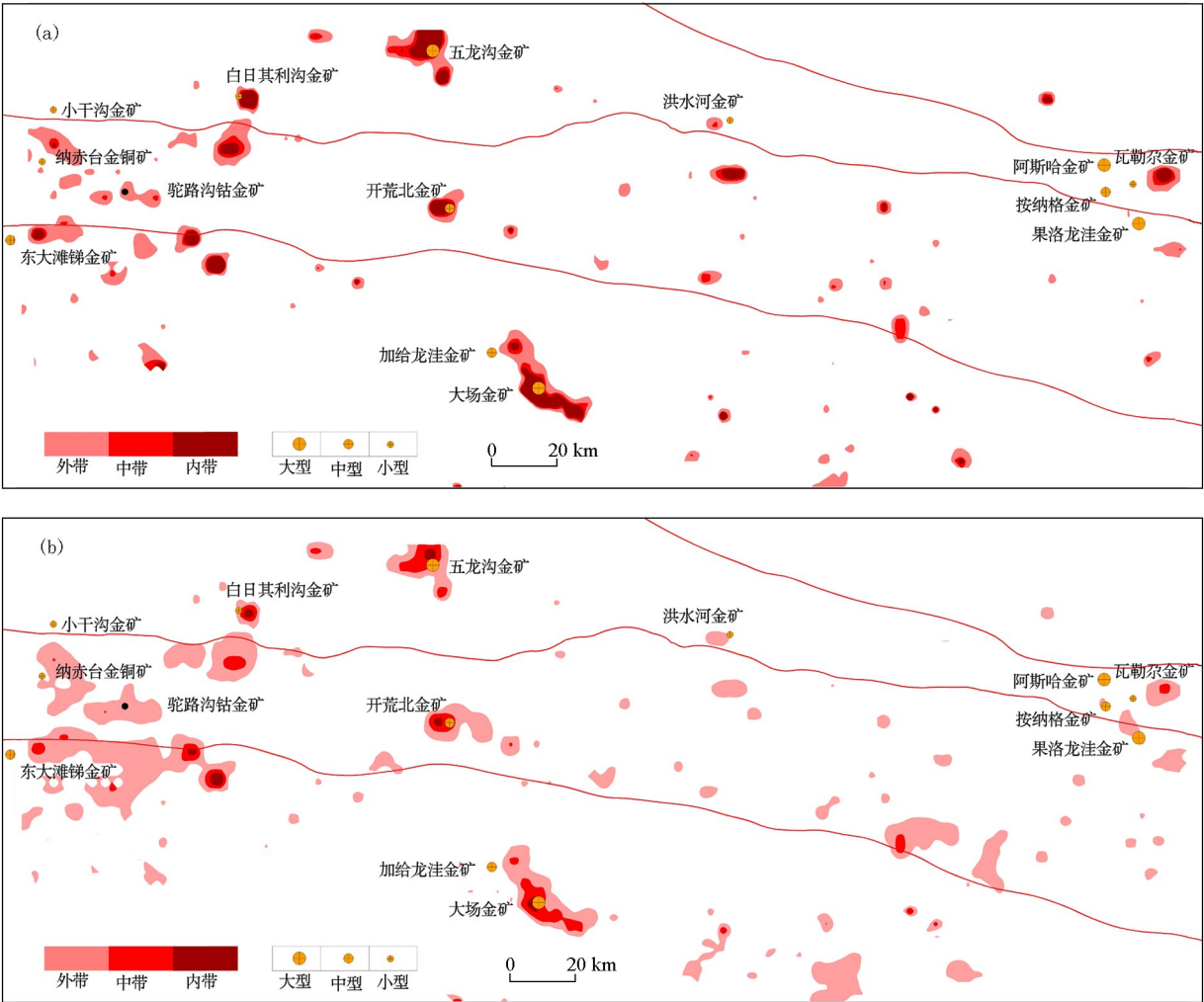
都处于 F_{3_2} 中; SiO_2 的最大值也处于 F_{3_2} 中; Fe_2O_3 、Mn 的最大值都处于 F_{4_2} 中,次小值都处于 F_{4_1} 中;该结果与因子载荷所反映的一致(表 3)。

根据本文分区标准化公式和 EDA 确定异常下限的方法,以 1、2、4 为异常的外、中、内带下限值圈定分区标准化异常图(图 6a);将剔除特高值后的原始数据平均值+2 倍标准离差作为异常下限,以 2、4 倍异常下限为中、内带圈定传统异常分布图(图 6b)。两种方法圈定的金异常点位分布对比如图 7 所示,异常点位特征对比列于表 5。

表 4 研究区各分区中常量组分中位数统计

Table 4 Major component median of different subdivision from the study area

指标	F1 ₁	F1 ₂	F2 ₁	F2 ₂	F3 ₁	F3 ₂	F4 ₁	F4 ₂
Mn	532.5	467.8	515	531	610	372.9	389	767
P	604	341	437	457.65	540.45	323	424	454
Ti	3746.5	1954.8	2704	3364.5	3481.55	2037	2609	3254.8
Zr	215	113	136	166	160	121	165	149.3
Al ₂ O ₃	11.6	10.8	9.1	13.6	12.1	7.5	9.6	11.94
CaO	4.4	4.83	9.1	2.32	6.2	2.64	4.7	3.3
Fe ₂ O ₃	4.68	3.06	3.6	4.84	4.8	2.77	3.06	5.25
K ₂ O	2.3	2.1	1.7	2.72	2.49	1.3	2	2.1
MgO	1.75	1.13	1.8	1.51	2.1	0.8	1.21	1.53
Na ₂ O	1.8	2	1.5	1.82	2.1	1.4	1.8	1.68
SiO ₂	64.6	69.28	59.39	66.945	60.215	78.2	69.69	67.71

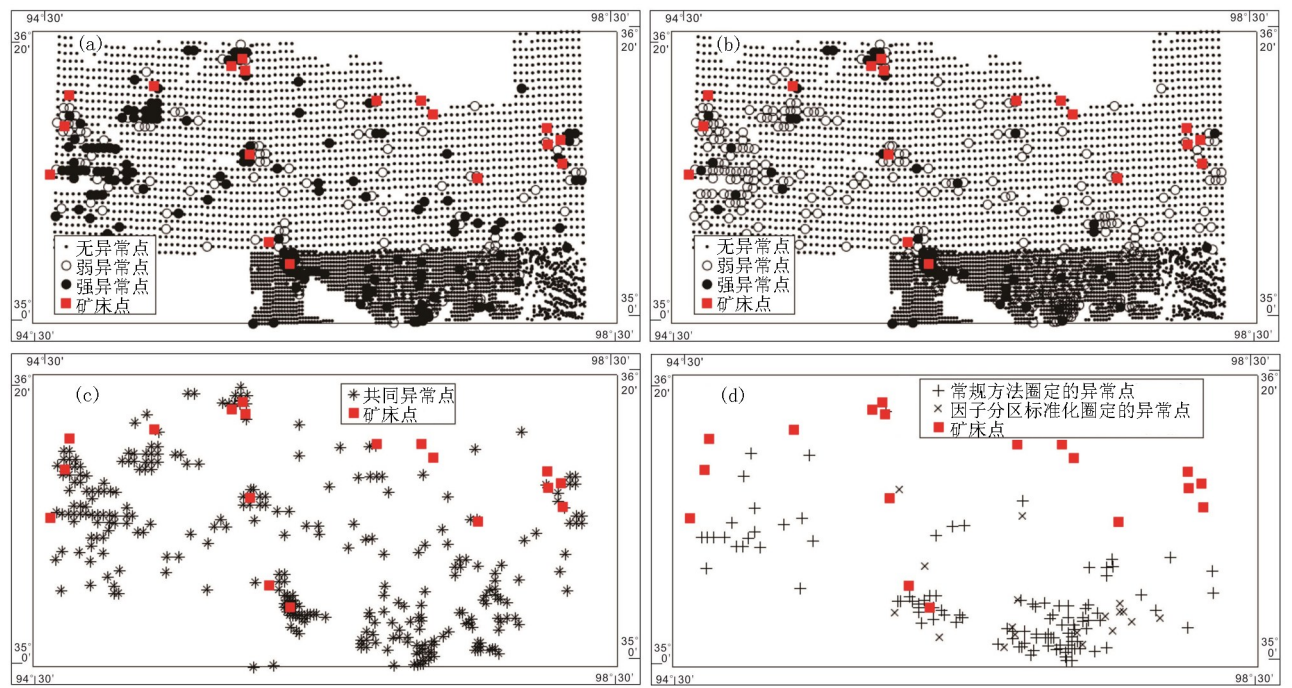


a—ILR 变换后因子分区标准化圈定的 Au 异常；b—传统化探数据方法圈定的 Au 异常

a—Au anomalies delineated from geochemical subdivision standardization of ILR;b—Au anomalies delineated from classical method

图 6 ILR 变换后数据因子分区标准化及传统化探数据处理方法圈定的 Au 异常分布对比

Fig.6 The comparison diagram of Au anomalies and deposits from geochemical subdivision standardized of ILR and classical method



a—因子分区标准化圈定的金异常点及金矿床点;b—传统方法圈定的金异常点及金矿床点;
c—两种方法都圈定出的金异常点及金矿床点;d—只在其中一种方法中圈定的金异常点
a—gold outliers detected from geochemical subdivised standardization;b—gold outliers detected from classical method;c—gold outliers detected from both method;d—gold outliers detected from one of two methods

图 7 传统方法与因子分区标准化方法圈定的金异常对比

Fig.7 Comparsion diagram of detecting Au outliers from subdivised standardization and classical method

表 5 两种方法圈定的金异常点位统计

Table 5 Point statistics of detecting Au outliers from two methods

处理方法	异常下限值	圈定的异常 点数	占样品总数 /%	两种都有的 异常点数	只在其中一种出现的 异常点数
传统方法	2.7×10^{-9}	415	10.4	307	108
因子分区标准化	1	325	8.1		18

从图 6 可知,用 IIR 变换后数据分区标准化圈定的异常和直接用原始数据圈定的异常在西部矿床(点)上都有良好的指示,而东北部指示效果不太理想,其原因推测与该区的表生地球化学环境有关,该区受风成砂影响严重。

直接用原始数据所圈定的异常较零星,浓集趋势相对较差,而用分区标准化所圈定的异常浓集趋势明显,与矿床点的对应关系也较好,因此用分区标准化圈定的异常更具有找矿指导意义。

结合图 7 和表 5 可知,用传统方法所圈定的金异常点数占总样品数的 10.4%,而用分区标准化所圈定的异常点数只占总样品数的 8.1%,两种方法都圈定出的异常点数占总样品数的 7.7%,且金矿床(点)几乎都处于两种方法都圈定出的异常点位处,表明分区标准化圈定的异常具有更高的可信度。此外,传统方法还在图幅的东南端圈定了大量的非矿异常点。东

南端处于北巴子区中,出露地层主要为三叠系巴颜喀拉群,其砂泥质类复理石沉积建造是 Au、Sb、As、Hg 的重要矿源层^[33]。麻多幅和扎陵湖幅地球化学调查结果显示,巴颜喀拉山群变砂岩中 Au 的背景含量达到 1.85×10^{-9} ^[34]。将本文采用的 4 001 件水系沉积物测量数据按昆北、昆中、昆南和北巴 4 个子区分别统计背景值,结果显示,Au 的背景值在昆南和北巴(昆南 1.59×10^{-9} ,北巴 1.54×10^{-9}) 远远高于昆北和昆中(昆北 0.76×10^{-9} ,昆中 0.96×10^{-9}),因此推断异常与该区的高背景有关。而分区标准化可以剔除高背景的影响,从而使异常圈定更合理。

4 结论

1) 地球化学数据是成分数据,由于其有闭合效应,因此在进行因子分析之前,要对数据进行变换。

从变换后的主成分分析来看,其结果更符合实际地质情况,更便于解释。

2) 因子分区标准化圈定的金异常和传统方法圈定的金异常对比认为,虽然两者在金矿床(点)上也都有显示,但是传统方法由于没有考虑地质条件的差异,在高背景地段也圈定了大量的非矿异常,而因子分区标准化考虑了地质条件的差异,因此圈定的异常更合理。

参考文献(References):

- [1] 李宝强,孙泽坤.区域地球化学异常信息提取方法研讨[J]. 西北地质, 2004,37(1):102-108.
Li B Q, Sun Z K. Study on the method of geochemical anomalies analysis[J]. Northwestern Geology, 2004,37(1)102-108.
- [2] 郝立波,李巍,陆继龙,等.确定岩性复杂区的地球化学背景与异常的方法[J]. 地质通报, 2007,26(12):1531-1535.
Hao L B, Li W, Lu J L, et al. Method for determining the geochemical background and anomalies in areas with complex lithology[J]. Geological Bulletin of China, 2007,26(12) 1531-1535.
- [3] 李长江,麻士华.矿产勘查中的分形、混沌与 ANN [M]. 北京:地质出版社,1999.
Li C J, Ma S H. Fractal, chaos and ANN in mineral exploration [M]. Beijing: Geological Publishing House, 1999.
- [4] 周蒂.分区背景校正法及其对化探异常圈定的意义[J]. 物探与化探,1986,10(4): 263-273.
Zhou D. Unit-wise adustment of geochemical background data and its significant geochemical anomaly delineation[J]. Geophysical and Geochemical Exploration, 1986, 10(4):263-273.
- [5] 史长义,张金华,黄笑梅.子区中位数阈值滤波法及弱小异常识别[J]. 物探与化探,1999,23(4):250-257.
Shi C Y, Zhang J H, Huang X M. Subregion median contrast filtering method and recognition of weak anomalies [J]. Geophysical and Geochemical Exploration, 1999,23(4):250-257.
- [6] 赵荣军.不同方法在柴川北部化探数据处理中的应用[J]. 地质与勘探,2006,42(3): 67-71.
Zhao R J. Application of different data processing method in geochemical exploration in the North Luanchuan [J]. Geology and Prospecting, 2006,42(3):67-71.
- [7] 刘大文.区域地球化学数据的归一化处理及应用[J]. 物探与化探, 2004,28(3): 273-275.
Liu D W. The normalization of regional geochemical data and its application[J]. Geophysical and Geochemical Exploration, 2004, 28(3):273-275.
- [8] 陈建国,夏庆霖.利用小波分析提取深层次物化探异常信息[J]. 地球科学:中国地质大学学报,1999,24(5):509-512.
Chen J G, Xia Q L. Wavelet-based extraction of geophysical and geochemical anomaly information[J]. Earth Science:Journal of University of Geosciences, 1999, 24(5):509-512.
- [9] 李随民,姚书振,韩玉丑.Sufer 软件中利用趋势面方法圈定化探异常[J]. 地质与勘探, 2007, 43(2):72-75.
Li S M, Yao S Z, Han Y C. Using tendency analysis method to deal with geochemical data based on the Suffer software[J]. Geology and Prospecting, 2007,43(2):72-75.
- [10] 陈希清,杨晓君,陈富文,等.应用 MAPGIS 数字高程模型提取区域地球化学异常信息的方法探讨[J]. 地球学报, 2009,30(1):119-125.
Chen X Q, Yang X J, Chen F W, et al. A discussion on the method for extracting regional geochemical anomaly based on mapgis digital elevation model [J]. Acta Geoscientica Sinica, 2009, 30(1):119-125.
- [11] 陈明,范继璋,矫希国.化探背景与异常划分中的 C 型转换法[J]. 长春地质学院学报,1996, 26(2): 227-230.
Chen M, Fan J Z, Jiao X G. The C-type transformation method for recognition of geochemical background and anomaly [J]. Journal of Chang Chun University of Earth Sciences, 1996,26(2):227-230.
- [12] 成秋明,赵鹏大,陈建国,等.奇异性理论在个旧锡铜矿产资源预测中的应用:成矿弱信息提取和复合信息分解[J]. 地球科学:中国地质大学学报, 2009,34(2):232-242.
Chen Q M, Zhao P D, Chen J G, et al. Application of singularity theory in prediction of tin and copper mineral deposits in Gejiu District, Yunnan, China: weak information extraction and mixing information decomposition[J]. Earth Science:Journal of University of Geosciences, 2009, 34(2):232-242.
- [13] 时艳香,纪宏金,陆继龙,等.水系沉积物地球化学分区的因子分析方法与应用[J]. 地质与勘探, 2004,40(5):73-76.
Shi Y X, Ji H J, Lu J L, et al. Factor analysis method and application of stream sediment geochemical partition [J]. Geology and Prospecting, 2004, 40(5):73-76.
- [14] 时艳香.区域地球化学单元的概念、方法及应用研究 [D]. 长春:吉林大学,2004.
Shi Y X. Research of regional geochemical Units: concept, methods and applications [D]. Changchun:Jilin University, 2004.
- [15] 赵玉岩,郝立波,陆继龙.利用水系沉积物识别基岩类型的方法研究——以大兴安岭浅覆盖区为例[J]. 吉林大学学报:自然科学版,2005,35(S):147-150.
Zhao Y Y, Hao L B, Lu J L. The technique of inveting bedrock types by residual stream sediment in Daxing'anling shallow overlay [J]. Journal of Jilin University: Earth Science Edition, 2005,35(S):147-150.
- [16] 郝立波,陆继龙,李龙,等.区域化探数据在浅覆盖区地质填图中的应用方法研究[J]. 中国地质, 2007,34(4):710-715.
Hao L B, Lu J L, Li L, et al. Method of using regional geochemical data in geological mapping in shallow overburden areas [J]. Geology in China, 2007, 34(4): 710-715.
- [17] 时艳香,郝立波,陆继龙,等.因子分类法在黑龙江塔河地区地质填图中的应用[J]. 吉林大学学报:自然科学版, 2008,38(5):899-903.
Shi Y X, Hao L B, Lu J L, et al. Application of factor classification in geological mapping in Tahe area, Heilongjiang Province [J]. Journal of Jilin University: Earth Science Edition, 2008, 38(5):899-903.
- [18] 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, 2008, 22(1):

78–82.

Dong Y. Discussion of applying factor analysis to the geochemical subareas measurement in stream sediment; A case study of Dulan area in Qinghai Province [J]. Mineral Resources and Geology, 2008, 22(1): 78–82.

[19] 赵少卿, 魏俊浩, 高翔, 等. 因子分析在地球化学分区中的应用; 以内蒙在石板井地区 1:5 万岩屑地球化学测量数据为例 [J]. 地质科技情报, 2012, 31(2): 27–34.

Zhao S Q, Wei J H, Gao X, et al. Factor analysis in the geochemical subdivisions: Taking 1:50 000 debris geochemical survey in the Shibanjing area of Inner Mongolia as an example [J]. Geological Science and Technology Information, 2012, 31(2): 27–34.

[20] Filzmoser P, Hron K, Reimann C. Univariate statistical analysis of environmental (compositional) data: problems and possibilities [J]. Science of the Total Environment, 2009, 407: 6100–6108.

[21] Pawlowsky-Glahn V, Buccianti A. compositional data analysis: theory and application [M]. London: John Wiley & Sons, 2011: 378.

[22] Aitchison J. The statistical analysis of compositional data [J]. Journal of the Royal Statistical Society. Series B (Methodological), 1982, 44(2): 139–177.

[23] Pawlowsky-Glahn V, Egozcue J J. Spatial analysis of compositional data: a historical review [J]. Journal of Geochemical Exploration, 2016, 164: 28–32.

[24] Pawlowsky-Glahn V, Egozcue J J, Tolosana-Delgado R. Modeling and analysis of compositional data [M]. UK: John Wiley and Sons, 2015: 272.

[25] Filzmoser P, Hron K, Reimann C. Principal component analysis for compositional data with outliers [J]. Environmetrics, 2009, 20(6): 621–632.

[26] Tucky J W. Exploratory data analysis [M]. Massachussetts: Addison-Wedley Reading, 1977: 506.

[27] Hoaglin D C, Mosteller F, Tukey J W. Understanding robust and exploratory data analysis (2nd edition) [M]. New York: John Wiley & Sons, 2000: 472.

[28] 史长义. 勘查数据分析 (EDA) 技术的应用 [J]. 地质与勘探, 1993, 11: 52–58.

Shi C Y. Application of exploratory data analysis technology [J]. Geolgy and Prospecting, 1993, 11: 52–58.

[29] Reimann C, Filzmoser P, Garrett R G. Background and threshold: critical comparison of methods determination [J]. Science of the Total Environment, 2005, 346(1–3): 1–16.

[30] Reimann C, Garrett R G. Geochemical background-concept and reality [J]. Science of the Total Environment 3, 2005, 50(1–3): 12–27.

[31] Grunsky E C. The evaluation of geochemical survey data: data analysis and statistical methods using geographic information systems [G] // Harris J R. Gis for the Earth Sciences, Geological Association of Canada Special Publication 44. Geological Association of Canada, St. John's, 2006: 229–283.

[32] Carranza E J M. Geochemical anomaly and mineral prospectivity mapping in GIS [M]. Oxford: Elsevier Scientific Publishing Company, 2009.

[33] 丁清峰. 东昆仑造山带成矿作用与矿产资源评价 [D]. 长春: 吉林大学, 2004.

Ding Q F. Metallogenesis and mineral resources assessment in Eastern Kunlun orogenic belt [D]. Changchun: Jilin University, 2004.

[34] 马进全, 郝平, 杨晓琴, 等. 中华人民共和国地球化学图说明书—I-47-[7] (麻多幅)、I-47-[8] (扎陵湖幅) [R]. 青海省地质调查院, 2000.

Ma J Q, Hao P, Yang X Q, et al. Specification of geochemical map of the People's Republic of China—I-47-[7], I-47-[8] [R]. Qinghai Geological survey Institute, 2000.

The application of ILR transformed data factor analysis to delineating geochemical anomalies

GENG Guo-Shuai^{1,2}, YANG Fan^{3,4}, GUO Jian-Na⁵

(1. School of Earth Sciences and Resources, China University of Geosciences, Beijing 100083, China; 2. Gold Geological Institute of CAPF, Langfang 065000, China; 3. Beijing Institute of Geology for Mineral Resources, Beijing 100012, China; 4. Research Center of Geochemical Survey and Assessment on Land Quality, China Geological Survey, Langfang 065000, China; 5. Natural Resources and Planning Bureau, Langfang 065000, China)

Abstract: The reliable detection of data outliers and unusual data behavior is one of the key task in the statistical analysis of applied geochemical data, and has remained a core problem. Factor analysis is a multivariate statistical analysis method, which is used to solve the problem of complex geological origin and superimposed mineralization; nevertheless, geochemical data are compositional data, there exist their closure effects, closure has a major influence on the covariance and correlation matrices, the very base of principal component analysis (PCA) and factor analysis (FA). So the authors applied isometric logratio-transformed (ILR) to ‘open’ the data before FA. The study area is located in the east of East Kunlun polymetallic mineralization zone. The authors used ILR transformed 11 major elements to conduct FA, extracted four public factors and calculated the four factor scores. According to the results of FA with EDA method, the authors standardized geochemical data and delineated Au anomaly. Compared with traditional method, this method can eliminate the influence of high background values.

Key words: regional geochemical exploration; ILR transformation; factor analysis; EDA method; anomalous delineation; East Kunlun