

研究报告

用主成份分析研究土壤中重金属含量与母质的关系

唐 诵 六

(中国科学院南京土壤研究所)

土壤中砷、汞、铅、铬、镉、铜、锌、钴、镍、锰等重金属元素,由于它们对生物和人体的潜在毒性,所以在环境保护工作中,它们经常是必要的监测项目。一般认为,这些元素在土壤中的背景值,主要受土壤类型的支配。然而,很多资料证明,这些微量重金属与土壤中的常量元素硅、铁、铝等不同,它们很可能不决定于现行的土壤分类,而在更大程度上受母质类型^[1,2]、颗粒机械组成^[3]等因素的制约。在一个工作地区,找出支配这些重金属元素分布的主要因素是十分必要的。因为这不仅有理论意义,还可以指导制订合理的工作方案诸如布点采样、统计、成图等,使工作量减少到最低程度并使实验结果更反映实际。本文目的是应用主成分分析法探讨影响不同土类、不同母质土壤中重金属元素含量分布的主导因素。

土壤样品和元素测定方法

35 个土壤样品采自南京郊区。它们都是地面被复度大、剖面发育较好的土壤。按土类有灰潮土、黄褐土和黄棕壤三大类,其中,灰潮土的母质为长江近代冲积物;黄褐土的母质为第四纪沉积的下蜀系粘土;黄棕壤是低山、丘陵上分布的微酸性土壤,其母质包括各种岩性不同的岩石,计有石灰岩、砂岩、页岩、花岗岩、玄武岩、辉长岩和橄榄辉长岩等。土壤过 100 目筛后用下述方法测定九种元素的含量:砷用银盐比色法,铬经氟化氢、

硫酸消解后用火焰原子吸收法测定,铅、铜、锌、钴、镍和锰用王水、高氯酸消解后再用火焰原子吸收法测定、汞用冷原子吸收法测定。

主成份分析

主成份分析,又称主组元分析,它是多元统计方法的一种。用主成份分析可以从大量样品、多个变量的原始观测数据中找出主要的决定变量大小的因素。它常被用来研究和解释地质现象。近年来,在土壤学的研究中,例如,水稻土成因分类^[4],红壤的属性分类^[5]土壤的波谱反射特性研究中^[6],主成份分析也得到了应用。

主成份分析的计算步骤是,首先根据样品的各变量测定值求出变量的相关系数矩阵。它是一个实对称矩阵。接着用雅可比法或克雷洛夫法(本文用雅可比法)求出该矩阵的特征值和特征向量。特征值的数目和变量数相等,每一特征值的特征向量数也与变量数相等。然后根据特征值的大小及其贡献率选取主成份。一般使累计贡献率达到 80% 即可。为了使主成份分析的结果能解释研究对象的问题,尚须进一步建立主成份方程式。将各样品的测定值代入方程式,可得出各样品的主成份值,并在图上标出位置。这样,根据各样品的位置,便可将样品分类,进而讨论研究对象受何种因素所支配。主成份分析的详细原理及计算步骤,已在有关的文献中详细介绍过^[4,7]。本文将用实例简单描述

表 1 南京 35 个样品 9 种元素的测定值

单位 ppm

编号	田间号	土壤	母 岩	As	Hg	Cr	Pb	Cu	Zn	Co	Ni	Mn
1	N17	灰潮土	长江冲积物	11.8	0.090	60	22.6	42.5	87.5	22.4	57.0	644
2	N19			14.7	0.118	83	31.2	48.5	98.7	23.9	60.4	732
3	N21			10.4	0.098	88	35.0	47.0	114	24.8	52.0	341
4	N5	黄刚土	下蜀黄土	7.3	0.026	71	16.2	26.6	68.6	17.6	32.0	594
5	N12			12.5	0.025	71	24.2	31.1	61.9	21.6	44.3	680
6	N16			9.4	0.025	60	14.9	25.8	58.8	19.1	34.4	555
7	N18			9.7	0.040	45	14.5	31.6	74.3	19.5	51.0	548
8	N20			9.9	0.045	60	17.6	39.3	77.5	18.1	49.7	517
9	N4	黄棕壤	崮山灰岩	202	1.243	44	125	37.6	244	28.2	38.9	911
10	N13		青龙灰岩	24.2	0.127	47	53.5	44.1	101	22.7	58.0	1391
11	N35			16.9	0.099	59	33.8	38.0	79.9	16.5	32.4	1081
12	N36			19.3	0.121	70	37.5	41.0	75.9	15.5	33.8	1285
13	N37			24.6	0.058	85	50.0	47.0	213	22.9	44.5	900
14	N38			24.6	0.123	60	86.7	43.0	152	15.7	42.0	1600
15	N10		赤山砂岩	2.3	0.019	18	6.8	10.6	20.2	6.3	15.5	118
16	N23		花 岗 岩	3.9	0.092	17	12.9	10.3	27.6	4.1	9.6	86.5
17	N39			2.6	0.023	9	7.5	9.5	8.0	3.0	4.5	65.0
18	N40			3.0	0.022	20	9.0	13.0	18.0	3.0	7.0	115
19	N41			2.9	0.022	11	4.5	6.5	9.0	1.0	2.5	30.0
20	N3		高家边页岩	9.4	0.043	50	21.5	17.0	52.6	13.4	21.2	170
21	N27		象山砂岩	11.8	0.058	44	44.5	25.0	68.0	13.4	22.6	643
22	N30			9.3	0.049	53	22.5	16.0	56.0	10.8	22.6	915
23	N31			7.1	0.030	47	19.5	11.0	30.0	4.7	10.4	143
24	N32			10.5	0.046	55	23.5	14.0	39.0	5.7	14.7	240
25	N33			7.2	0.068	34	9.0	12.0	33.0	6.3	10.0	248
26	N1		黄马青砂页岩	5.0	0.054	59	27.2	41.9	82.4	14.3	31.7	732
27	N28			10.2	0.028	69	28.0	64.0	105	18.6	28.8	1010
28	N29			8.0	0.024	71	33.8	15.0	69.0	17.2	29.8	620
29	N42			8.7	0.080	76	29.1	11.0	63.5	11.4	25.0	585
30	N43			7.2	0.079	70	28.1	10.0	53.8	10.5	23.5	360
31	N44			7.2	0.059	60	10.5	24.0	40.2	10.4	18.8	545
32	N9		玄 武 岩	6.2	0.099	130	15.7	52.6	114	58.1	93.1	919
33	N45		辉 长 岩	5.2	0.076	127	18.8	154	163	60.0	61.0	1458
34	N46			8.2	0.022	76	15.2	176	171	44.0	47.0	1020
35	N47		橄榄辉长岩	4.2	0.106	548	16.1	185	110	97.0	175	1925

注：编号 1—3 是灰潮土，4—8 是黄刚土，9—35 是黄棕壤，1—3 是长江冲积物，4—8 是下蜀黄土，10—14 是青龙灰岩，16—19 是花岗岩，21—25 是象山砂岩，26—31 是黄马青砂页岩，33—34 是辉长岩。

表 2 九个变量的统计量

变量 统计量	As	Hg	Cr	Pb	Cu	Zn	Co	Ni	Mn
均 值	15.4	0.095	73	27.6	40.6	81.2	20.0	37.3	678
标 准 差	33.0	0.203	87	23.3	43.6	54.6	19.0	31.1	470
变异系数	2.14	2.14	1.19	0.84	1.07	0.67	0.95	0.83	0.69

表 3 九个变量的相关系数矩阵

	As	Hg	Pb	Zn	Mn	Co	Ni	Cr	Cu
As		+++	+++	+++					
Hg	0.9854		+++	+++					
Pb	0.8192	0.7796		+++	+				
Zn	0.5955	0.5654	0.6996		+++	+++	++		+++
Mn	0.1759	0.1798	0.4285	0.6521		+++	+++	+++	+++
Co	0.0696	0.1311	0.0585	0.5506	0.7121		+++	+++	+++
Ni	0.0307	0.0884	0.0903	0.4680	0.7079	0.9297		+++	+++
Cr	-0.0708	-0.0026	-0.0466	0.2584	0.5977	0.8509	0.8859		+++
Cu	-0.0115	0.0242	0.0036	0.5683	0.6979	0.8690	0.7307	0.6933	

+++ $P < 0.001$ ++ $P < 0.01$ + $P < 0.1$

上述计算。而本文的全部计算工作系由一台日制 SHARP PC-1500 袖珍计算机完成*。

结 果 与 讨 论

南京地区 35 个土壤样品所属土类、母质以及九种元素的测定值列于表 1。从表 1 及表 2 的数据可以看出,不同样品中各元素的含量变化很大,尚不足以判断究竟是土壤类型还是母质类型控制着各元素的含量。表 3 是 35 个样品九种元素间的相关系数矩阵。矩阵中的数值是相关系数, + 号表示相关系数的置信度。从表 3 可以看出,元素钴、镍、铬、铜、锰、锌之间有着极显著相关,元素砷、汞、铅、锌之间也有极显著相关。然而,在钴、镍、铬、铜、锰与砷、汞、铅之间却没有相关性。表现在表的右上部有一个空白区。这表明作为铁族元素的钴、镍、铬、铜、锰是一个集团,而元素砷、汞、铅可视为另一个集团。

用雅可比法求得该矩阵的九个特征值。由表 4 可见,头两个特征值的累计贡献率已

表 4 九个特征值及其贡献率

特 征 值	贡 献 率 %	累计贡献率 %	主 成 份
4.668	51.86	51.86	1
2.953	32.81	84.67	2
0.611	6.79	91.46	
0.404	4.48	95.94	
0.178	1.98	97.92	
0.107	1.19	99.11	
0.043	0.48	99.59	
0.029	0.32	99.91	
0.008	0.09	100.00	

达到 85%。因此,就选取它们作为两个主成份。至此,35 个样品、9 个变量测定值的全部信息的 85% 已被归纳为两个主成份。求得该两主成份的特征向量(表 5)。它们分别代表 9 个变量在该主成份中的权系数值。由表 5 可见,第一主成份以钴为代表,主要反映元素钴、镍、锰、铜、锌、铬的富集程度。第二

* 唐诵六,适用于 SHARP PC-1500 机的主成份分析 Basic 语言算法程序、资料,1984。

表 5 两个主成分的特征向量

主 成 份	特 征 向 量								
	As	Hg	Pb	Zn	Mn	Co	Ni	Cr	Cu
1	0.1813	0.1959	0.2104	0.3661	0.3941	0.4162	0.3987	0.3477	0.3764
2	0.5119	0.4851	0.4730	0.2373	0.0472	-0.2104	-0.2189	-0.2825	-0.2198

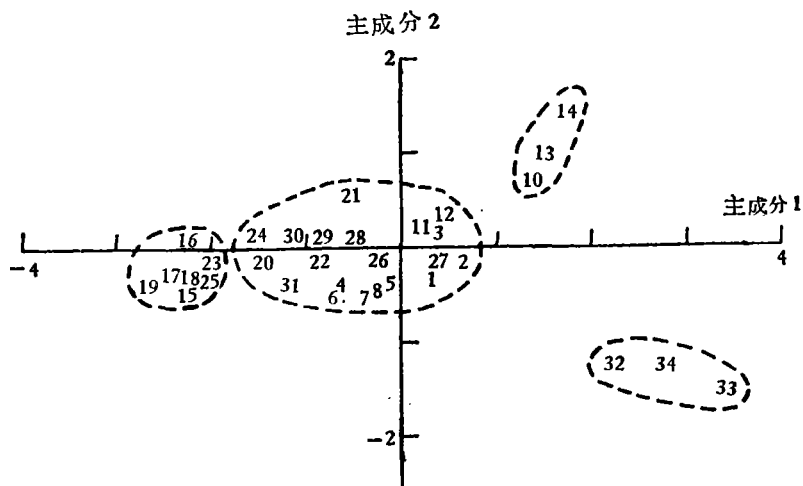


图 1 35 个样品按第 1, 2 主成份分类

主成分以砷为代表,主要反映元素砷、汞、铅的富集程度,与之相斥的元素则有钴、镍、铬、铜。

进一步求得该两个主成分的主成份方程:

$$\begin{aligned}
 Y_1 = & 0.005\text{As} + 0.967\text{Hg} + 0.004\text{Cr} \\
 & + 0.009\text{Pb} + 0.009\text{Cu} \\
 & + 0.007\text{Zn} + 0.002\text{Co} \\
 & + 0.013\text{Ni} + 0.001\text{Mn} \\
 & - 3.098 \quad (1)
 \end{aligned}$$

$$\begin{aligned}
 Y_2 = & 0.015\text{As} + 2.394\text{Hg} - 0.003\text{Cr} \\
 & + 0.02\text{Pb} - 0.005\text{Cu} \\
 & + 0.004\text{Zn} - 0.011\text{Co} \\
 & - 0.007\text{Ni} - 0.0002\text{Mn} \\
 & - 0.385 \quad (2)
 \end{aligned}$$

式中,各元素的系数值保留至小数后三位,但在实际计算时则为小数后九位。利用公式

(1)及(2),分别计算出 35 个样品的两个主成分值(表 6)。最后,以主成份 1 为横坐标,以主成份 2 为纵坐标,将 35 个样点逐个点入座标图中。将相互靠近的点用围线勾出,便得到图 1。

根据图 1 中样点的位置,大致可将 35 个样品分成四组。第一组是位于第一象限内离原点较远的样品 9、10、13、14(样品 9 越出图外未标出)。第二组是位于第四象限内离原点较远的样品 32、33、34、35(样品 35 越出图外未标出)。第三组是原点附近的 20 个样点。第四组是第一主成份值小于 -2 的 7 个样点。根据图 1 及表 1,可以作如下的讨论。从土壤类型看,第一、二、四组均为黄棕壤,第三组内则兼有黄棕壤、黄褐土及灰潮土三类。由此可见,根据九种重金属元素含量所得出的分组并不遵从土壤类型。从母质类

表 6 各土壤的主成份值

编 号	Y1	Y2	编 号	Y1	Y2
1	0.214	-0.272	19	-2.781	-0.257
2	0.704	-0.091	20	-1.403	-0.035
3	0.389	0.017	21	-0.605	0.493
4	-0.619	-0.428	22	-0.777	-0.048
5	-0.207	-0.379	23	-2.014	-0.025
6	-0.705	-0.457	24	-1.667	0.094
7	-0.383	-0.458	25	-1.977	-0.121
8	-0.275	-0.426	26	-0.289	-0.127
9	4.359	8.307	27	0.394	-0.143
10	1.294	0.643	28	-0.571	0.012
11	0.179	0.243	29	-0.825	0.143
12	0.457	0.321	30	-1.168	0.123
13	1.545	0.903	31	-1.264	-0.363
14	1.793	1.667	32	2.152	-1.318
15	-2.271	-0.381	33	3.425	-1.463
16	-2.244	0.048	34	2.498	-1.209
17	-2.645	-0.252	35	7.688	-4.476
18	-2.415	-0.250			

型看,第一组为石灰岩土壤,第二组为基性火成岩(玄武岩、辉长岩、橄榄辉长岩)土壤,第四组为酸性岩(花岗岩、部分砂岩)土壤,第三组则为各种中性及中酸性岩类,包括沉积物及沉积岩类。仅有两个石灰岩土壤(11, 12)例外,它们夹在第三组内,然而仍处于第一象限中。由此可见,分组基本上反映了不同的母质类型。从地球化学的角度看,第一主成份主要反映了第一过渡系元素的富集程度,而第二主成份则反映部分亲硫元素的富集程度。不同岩类在第一主成分中的排列次序为:基性火成岩 > 石灰岩 > 中酸性岩 > 酸性火成岩。在第二主成份中的次序为:石灰岩 > 中酸性岩 > 基性火成岩。以上这些都与微量重金属元素在岩石、矿物中的分布规律相吻合。

结 语

根据以上的讨论可见,主成份分析有助

于阐明土壤中微量重金属元素的分布规律。就南京地区而言,是成土母质类型对土壤中微量重金属元素的分布起着主导作用。现行的土壤分类不一定能反映这些元素的分布规律。这提示,在进行土壤环境中重金属元素的调查时,在设计布点采样方案、确定统计和制图单元、对测定数据进行分析判断等环节中,都应对母质状况给予更多的重视。

参 考 文 献

- [1] Ure, A. M. et al., *Geoderma*, 22:1 (1979).
- [2] 杨学义, 环境中若干元素的自然背景值及其研究方法, 16—20 页, 科学出版社, 1982 年.
- [3] 沈碧珍等, 土壤学报, 20: 440—444 (1983).
- [4] 刘多森, 土壤学报, 16: 172—183 (1979).
- [5] 王明珠, 土壤, 15: 59—64 (1983).
- [6] 徐彬彬, 土壤学报, 18: 176—184 (1981).
- [7] Davis, J. C., *Statistics and Data Analysis in Geology*. John Wiley and Sons Inc., 1973.