第13卷 第1期　　海 洋 与 湖 沼　　Vol. 13, No. 1
1982年1月　　OCEANOLOGIA ET LIMNOLOGIA SINICA　　Jan., 1982

A-B 相关法中的 A,B 端元*

范 守 志
（中国科学院海洋研究所）

一、问题的提出

在运用Q因子分类法[1]对一批样品进行统计分群时，需要首先选定某样品作为参比样品，即A样品。

在所研究的区域的学科背景比较明朗的情况下，例如在有河流注入的封闭或半封闭的沉积盆地中，或者在深水区和浅水区并存的海区中，A样品可由有关的学科特征来选定[2]。但在背景复杂的海区，以及在研究古沉积环境时，A样品如何产生呢？虽然这时可以利用各种已知的现代沉积环境所提供的典型样品，逐一充当A样品，进行统计对比，然而从数学地质研究的角度来看，寻求一种由所研究的那批样品本身确定A样品的方法无疑是十分重要的。本文的目的即在于此。

另一方面，众所周知，如果某区域的学科背景十分复杂，那必定是因为在这一区域中有着若干对矛盾着的环境因素共同起作用的结果。但无论如何，其中总有一对矛盾或多或少地起着主导作用；从而，一个海区的环境划分总是首先可分成两大基本类型，以及一个可能有的过渡型。于是，在遍布全区采集的一批样品中，也总是存在这样一对样品A和B，分别相应于这两种基本类型，它们在矿物组成或其它的学科组成方面是富有特征的；其余各样品的组成状态则相当于A和B依各种不同的比例混和时的产物。当然，这种混和状态可能相应于某种实际存在的混和过程，也可能是由于这些样品的生成环境介于两种典型环境之间的结果。

基于上述观点，本文提出了一种由所研究的样品中确定A样品和B样品的数学方法，并易于在电子计算机上求一解。

文中还例证了方法的有效性。

二、数学分析

设在某海区采集了N块标本，其中A和B是典型标本，其余$N-2$块在组成上是A和B依各种比例混和的产物。

兹证明，这时可依如下的法则从这N块样品中挑出A和B来：A和B之间的Q值是这N块样品中的任意一对样品之间的Q值中的最小者，即

$$Q_{A,B} = \mathrm{Min}\, Q_{i,j};$$
$$i, j = 1, 2, 3, \cdots, N.$$

* 中国科学院海洋研究所调查研究报告第651号。
本编辑部收到稿件日期：1979年1月3日。

为此我们需要证明下述定理 1 和定理 2。

定理 1　如果样品 C 是由样品 A 和 B 依某种任意的比例混和而成的，那么总有

$$Q_{A,C} \geqslant Q_{A,B},$$

及

$$Q_{B,C} \geqslant Q_{A,B},$$

且等号仅当

$$Q_{A,B} = 1 \text{ 时成立。}$$

定理 2　如果样品 C 和 D 都是由 A 和 B 混和而成的，那么不管混和比例相同与否，总有

$$Q_{C,D} \geqslant Q_{A,B},$$

且等号仅当

$$Q_{A,B} = 1 \text{ 时成立。}$$

定理 1 的证明

各取 M 克的 A 样品和 B 样品，对所研究的 n 种成份的含量进行测定，结果得到两个 n 维的组合向量[1]:

$$\boldsymbol{A} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{n-1} \\ A_n \end{pmatrix}; \quad \boldsymbol{B} = \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_{n-1} \\ B_n \end{pmatrix}. \tag{2-1}$$

设 M 克的 C 样品相当于 m_A 克的 A 样品和 $M - m_A$ 克的 B 样品的混和物，其组合向量记为 $\boldsymbol{C}$. 于是

$$\boldsymbol{C} = \frac{m_A}{M}\boldsymbol{A} + \frac{M - m_A}{M}\boldsymbol{B}. \tag{2-2}$$

用 A 和 B 表示向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的长度，即

$$A = \sqrt{\sum_{K=1}^{n} A_K^2}, \quad B = \sqrt{\sum_{K=1}^{n} B_K^2}; \tag{2-3}$$

则

$$\boldsymbol{A} = A\boldsymbol{a}, \quad \boldsymbol{B} = B\boldsymbol{b},$$

这里，$\boldsymbol{a}$ 和 $\boldsymbol{b}$ 是归一化的组合向量。

于是

$$\left.\begin{aligned} &\boldsymbol{C} = k_A A\boldsymbol{a} + k_B B\boldsymbol{b}, \\ &k_A = m_A/M, \quad k_B = 1 - k_A, \end{aligned}\right\} \tag{2-4}$$

其中

且

$$0 \leqslant k_A, \ k_B \leqslant 1.$$

在此 n 维欧氏空间中实行正交转轴变换，变换方阵是 $\boldsymbol{R}$,

$$\boldsymbol{a} = \boldsymbol{R}\boldsymbol{a}', \ \boldsymbol{b} = \boldsymbol{R}\boldsymbol{b}', \ \boldsymbol{C} = \boldsymbol{R}\boldsymbol{C}',$$

则

$$Q_{A,B} = \boldsymbol{a}' \cdot \boldsymbol{b}',$$

$$Q_{A,C} = \boldsymbol{a}' \cdot \boldsymbol{c}' = \frac{\boldsymbol{a}' \cdot \boldsymbol{C}'}{C'}, \tag{2-5}$$

这里 C' 是 $\boldsymbol{C}'$ 即 $\boldsymbol{C}$ 的长度。

业已证明[3],存在这样的 $\boldsymbol{R}$ 变换,使得

$$\boldsymbol{a}' = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}; \quad \boldsymbol{b}' = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \sqrt{1 - Q_{A,B}^2} \\ Q_{A,B} \end{pmatrix}, \tag{2-6}$$

于是

$$\boldsymbol{C}' = \boldsymbol{R}^*\boldsymbol{C} = \boldsymbol{R}^*(k_A A\boldsymbol{a} + k_B B\boldsymbol{b})$$
$$= k_A A\boldsymbol{a}' + k_B B\boldsymbol{b}',$$

即

$$\boldsymbol{C}' = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ k_B B\sqrt{1 - Q_{A,B}^2} \\ k_A A + k_B B Q_{A,B} \end{pmatrix}, \tag{2-7}$$

长度

$$C' = \sqrt{(k_A A)^2 + (k_B B)^2 + 2k_A A k_B B Q_{A,B}}, \tag{2-8}$$

因此

$$\boldsymbol{a}' \cdot \boldsymbol{C}' = k_A A + k_B B Q_{A,B}, \tag{2-9}$$

把 (2-8) 及 (2-9) 代入 (2-5),

$$Q_{A,C} = \frac{k_A A + k_B B Q_{A,B}}{\sqrt{(k_A A)^2 + (k_B B)^2 + 2k_A A k_B B Q_{A,B}}}. \tag{2-10}$$

注意

$$0 \leqslant Q_{A,B} \leqslant 1,$$

在 (2-10) 的分母中用 1 置换 $Q_{A,B}$, 就有

$$Q_{A,C} \geqslant \frac{k_A A + k_B B Q_{A,B}}{k_A A + k_B B}, \tag{2-11}$$

(等号仅当 $Q_{A,B} = 1$ 时成立。)

此式即

$$Q_{A,C} \geqslant Q_{A,B} \times \frac{\dfrac{k_A A}{Q_{A,B}} + k_B B}{k_A A + k_B B};$$

但

$$\frac{k_A A}{Q_{A,B}} \geqslant k_A A,$$

且仍旧仅当 $Q_{A,B} = 1$ 时其中的等号才成立,故有

$$Q_{A,C} \geqslant Q_{A,B}, \tag{2-12}$$

且等号仅当 $Q_{A,B} = 1$ 时才成立。

类似地,可证

$$Q_{B,C} \geqslant Q_{A,B}.$$

证毕。

定理 2 的证明

现在我们可写出

$$\boldsymbol{C} = k_A\boldsymbol{A} + k_B\boldsymbol{B},$$
$$\boldsymbol{D} = l_A\boldsymbol{A} + l_B\boldsymbol{B}.$$

仍然实行所述 $\boldsymbol{R}$ 变换，得到

$$\boldsymbol{C}' = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ k_B B\sqrt{1 - Q^2_{A,B}} \\ k_A A + k_B B Q_{A,B} \end{pmatrix}, \quad \boldsymbol{D}' = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ l_B B\sqrt{1 - Q^2_{A,B}} \\ l_A A + l_B B Q_{A,B} \end{pmatrix}.$$

于是

$$Q_{C,D} = \frac{\boldsymbol{C}' \cdot \boldsymbol{D}'}{C'D'}.$$

而

$$\boldsymbol{C}' \cdot \boldsymbol{D}' = k_A A l_A A + k_B B l_B B + [k_A A l_B B + k_B B l_A A]Q_{A,B};$$
$$C' = \sqrt{(k_A A)^2 + (k_B B)^2 + 2k_A A k_B B Q_{A,B}},$$
$$D' = \sqrt{(l_A A)^2 + (l_B B)^2 + 2l_A A l_B B Q_{A,B}}.$$

由于

$$Q_{A,B} \leqslant 1,$$

故

$$C' \leqslant k_A A + k_B B, \quad D' \leqslant l_A A + l_B B,$$

而

$$Q_{C,D} \geqslant \frac{k_A A l_A A + k_B B l_B B + [k_A A l_B B + k_B B l_A A]Q_{A,B}}{k_A A l_A A + k_B B l_B B + k_A A l_B B + k_B B l_A A};$$

类似于由 (2–11) 导出 (2–12) 的各步不等代换，由此不难证明

$$Q_{C,D} \geqslant Q_{A,B},$$

且等号仅当 $Q_{A,B} = 1$ 时成立。

证毕。

综上所述，$Q_{A,B}$ 是一切 Q 因子中的最小者。

现在，结合文章[1]和[3]中的结论与上述法则，可以把 $A–B$ 相关法的各步骤系统归纳如下：

1. 用电子计算机算出所有的 $Q_{i,j}$; $i, j = 1, 2, \cdots N$;
2. 用电子计算机依 Q 值最小法则确定 A 样品及 B 样品；
3. 利用算得的 $Q_{A,j}$ 及 $Q_{B,j}$ 点绘 $A–B$ 相关图；
4. 在 $A–B$ 相关图上作出所要的信任椭圆，并分类；
5. 在站位图上综合其它有关资料进行成因解释。

值得指出的是，对于分出的各类样品，也可以重复上述步骤分出可能有的亚群。

三、 例

即以文章[1]中引用的 27 块白云岩标本的 13 种化学成份的测定资料为例。在那里，$\boldsymbol{A}$

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	1.000																										
2	998	1.000																									
3	992	997	1.000																								
4	977	986	982	1.000																							
5	976	988	987	993	1.000																						
6	979	990	989	992	999	1.000																					
7	937	954	950	989	980	975	1.000																				
8	984	985	981	971	966	965	937	1.000																			
9	998	998	991	984	983	986	948	984	1.000																		
10	989	994	998	973	982	983	939	980	986	1.000																	
11	995	998	993	991	991	993	963	980	998	987	1.000																
12	965	980	980	991	999	997	984	959	974	974	985	1.000															
13	999	997	992	974	972	976	932	985	997	989	993	961	1.000														
14	999	997	993	978	977	979	938	992	997	990	994	967	999	1.000													
15	994	996	996	977	978	980	941	994	992	995	990	970	995	998	1.000												
16	889	916	918	953	967	962	974	881	907	909	928	976	882	890	897	1.000											
17	844	876	877	930	939	931	966	851	866	867	890	953	836	849	858	992	1.000										
18	731	774	781	847	861	849	907	745	757	773	790	883	722	738	755	958	981	1.000									
19	478	533	544	637	655	640	732	482	512	532	558	686	466	481	501	824	866	942	1.000								
20	894	917	912	957	964	960	976	895	915	900	932	973	885	898	899	990	988	942	794	1.000							
21	870	898	897	950	953	946	981	871	889	886	912	964	862	872	879	994	997	968	843	990	1.000						
22	842	873	874	921	938	932	953	841	866	863	890	951	833	846	852	993	996	973	860	991	990	1.000					
23	817	849	851	901	916	910	939	816	842	839	866	928	809	821	827	974	982	961	859	968	976	982	1.000				
24	795	832	837	888	908	900	931	789	819	828	848	924	786	797	808	984	990	988	907	968	982	992	979	1.000			
25	708	750	754	833	841	829	898	713	736	740	772	863	697	711	723	948	973	994	954	936	964	967	957	983	1.000		
26	540	593	606	683	710	696	767	544	572	597	615	739	528	544	564	864	899	962	993	835	874	898	891	938	966	1.000	
27	456	511	524	619	633	618	718	458	488	510	537	664	444	458	478	805	849	929	997	772	828	840	841	891	945	985	1.000

样品是第 27 号样品；这是借用兄弟单位的同志用复杂得多的方法(主因素分析法 Q 型)得出的结果。现在，我们用本文所述的最小 Q 因子法则，在国产的 DJS-6 型电子计算机上重新计算，得到的结果的确是一样的：

$$Q_{27,13} = \text{Min}\, Q_{i,j};$$
$$i, j = 1, 2, \cdots, 27.$$

打印输出的各个 Q 因子见上页表。

至于 Q 因子分类法的适用性，当然，还有待于进一步的实践来检验。值得指出的是，已用该方法对东海 177 个表层沉积物样品中的 122 种底栖有孔虫的组合状态进行了区系划分，结果是满意的[1)]。

参 考 文 献

[1] 范守志，1979。应用Q因子进行样品分类的 A-B 相关法。海洋与湖沼 **10** (4):319—328。
[2] 陈丽蓉、范守志，1981。渤海沉积物中矿物组合的统计分析。海洋与湖沼 **12** (3):235—239。
[3] 范守志，1981。A-B 相关法中的信任椭圆。海洋与湖沼论文集 1981：123—136。

THE A AND B END-MEMBER SAMPLES IN THE A-B CORRELATION METHOD*

Fan Shouzhi
(Institute of Oceanology, Academia Sinica)

Abstract

In this paper, we discuss how to select the A and B end-members in the A-B correlation method directly from the data matrix of samples studied.

Provided $\boldsymbol{X}_i$ (any one of samples studied) can be considered as a compound vector

$$\boldsymbol{X}_i = a_i\boldsymbol{A} + b_i\boldsymbol{B}$$
$$(i = 1,\ 2,\ \cdots N)$$

where a_i and b_i are two positive numbers or zero and N is the total number of samples, the following formula is always correct

$$Q_{A,B} \leqslant Q_{i,j}$$
$$(i,\ j = 1,\ 2,\ \cdots N)$$

where $Q_{A,B}$ and $Q_{i,j}$ are the values of factor Q between end-members A and B and between any two samples $\mathbf{X}_i$ and $\mathbf{X}_j$, separately.

According to the formula above, we can select the end-members A and B by means of computer based on

$$Q_{A,B} = \text{Min}\, Q_{i,j}$$
$$(i,\ j = 1,\ 2,\ \cdots N).$$

1) 将另有专文介绍。

* Contribution No. 651 from the Institute of Oceanology, Academia Sinica.