中国地质科学院

矿床地质研究所所刊

1986年　第1号

（总第17号）

地质出版社

中国地质科学院

矿床地质研究所所刊

1986年第1号（总第17号）

（数学地质专刊）

地 质 出 版 社

内 容 简 介

本刊为数学地质专刊，共十七篇文章。它们涉及数学地质方法，方法的地质应用，计算机在地质中的应用和计算机绘图。具体内容包括如下：近年来发展起来的新方法，如地质总体分解，地质数据预处理方法，相合分析，秩评定，一种利用混合型数据评价含矿性的方法，定和数据Q型因子分析等等；上述方法在矿产资源评价、地球化学等领域中的应用；矿产资源评价中的数学模型和勘探地质学中的数学模型综述；计算机专家系统在研究南岭花岗岩类含矿性方面的应用；计算机绘图（谱系图、三角图解、直方图、频率分布曲线、坐标散点图和玫瑰图等）；STYR程序库和BMDP程序库简介。

中国地质科学院
矿床地质研究所所刊
1986年 第1号
（总第17号）

●

责任编辑：张肇新 雍拥瑜 张中民 倪瑞兰
地质出版社出版
（北京西四）
北京昌平沙河建华印刷厂印刷
（北京海淀区学院路29号）
新华书店北京发行所发行·各地新华书店经售

●

开本：787×1092 1/16 印张：13 3/4 字数323,000
1986年10月北京第一版·1986年10月北京第一次印刷
印数：2500册 · 定价：3.30元
统一书号：13038·新283

中国地质科学院
矿床地质研究所所刊

1986年 第1号
(总第17号)

目　录

中国地质科学院 矿床地质研究所所刊
BULLETIN OF THE INSTITUTE OF MINFRAL DEPOSITS
1986年 CHINESE ACADEMY OF GEOLOGICAL SCIENCES 第1号

地质总体分解

李裕伟 李纯杰
（矿床地质研究所）

引 言

近年来，地质学中的“混合”问题日益引起人们的注意。这一问题的解决将为地质观测数据的成因解释另辟蹊径，特别是有助于识别那些隐含的矿化现象。

从数学地质的观点而言，存在两类混合问题：第一类混合问题是把任何一个观测样品看成是多个“理想样品”的按比例混合。所谓理想样品，可视为一个相对独立的地质过程的代表；而每个观测样品，则多半是若干个地质过程叠加的产物。这类混合问题是Q型因子分析研究的对象。第二类混合问题是把每个观测样品看成纯是某一地质过程的产物。我们经常碰到这样的观测样品的集合，它是代表不同地质过程的样品的混合，而这正是本文所要讨论的混合问题。下面所提到的混合，均属于第二类混合，它可以通过混合总体的分解来求得解决。

在地质学研究中，人们总是力图将观测样品按一定的准则分类，以使某一类样品代表某种独立的地质成因、环境或条件。当问题处于宏观分析阶段时，进行这种样品的成因分类不会出现多大困难。例如，对于一组取自花岗岩类的岩石样品，可以按花岗岩、花岗闪长岩、二长岩、正长岩来予以分类。但是，当问题处于亚宏观或微观分析阶段时，人们立即发现每一类又可能是若干种隐含的亚类的混合，要识别是否存在这种隐含的混合及把样品进一步按隐含的类型来划分，仅仅依靠野外观察和岩矿手段就不够了。

总体分解提供了一条识别混合与分离混合的途径。我们假定，一个相对独立的地质过程可用一种统计分布来静态地予以描述，那么观测子样的统计分布就有可能是代表不同地质过程的统计分布的混合。于是，混合的识别和分离问题就化为将一个观测的统计总体分解为若干个理论子总体问题。如果这种分解获得成功，就证实了在观测子样中蕴含着多个地质成因过程的信息，并可进一步将样品按所代表的成因过程划分为亚类。

早在1894年，皮尔逊（Pearson.K）[21]就提出了用矩量法解一维正态总体分解问题；哈丁（Harding,J.P.,1949）[16]很早就用图解法分离一维正态混合总体；辛克莱（Sinclair,A.J.,1981）[1]则是在地质学中用图解法进行一维正态或对数正态总体分解的大力倡导者；麦克卡门（McCammon R B.,1969）[20]、克拉克和加里特（Clark,I. and Garnett.R H T.,1974）[7]、克拉克（1977）[8]利用非线性最小二乘法求解一维正态总体分解问题；另一个克拉克（Clark,M W.,1977）[9]则详细讨论了矩量法一维正态总体分解；德依（Day,N.E.,1969）[14]提出了多维正态总体分解法；李裕伟（1984，

1985）[18,19]在研究空间图形识别问题时，提出了一种双变量约束总体分解方法；里德（Rider,P.R.,1961）[23]应用矩量法实现了指数总体分解；哈色尔布拉德（Hasselblad, V.,1969）[19]推导出指数总体分解的最大似然方程；布里斯克（Blischke,1962,1963, 1964）[4,5,6]和科享（Cohen,A.C.,1963,1965,1966）[10,11,12,]对发展离散型总体——二项分布、泊松分布的分解方法作出了重要的贡献。迄今为止，在地质学中已拥有了一套从图解到解析，从一维到多维，从连续型到离散型，从无约束到有约束的总体分解方法、程序和应用经验，这为使用总体分解技术解决各种地质问题奠定了一个良好的基础。

地质学中的总体分解问题非常广泛，例如，分离地球化学背景与异常，识别矿化阶段，划分沉积相，判定岩浆岩的期次，圈定古生物的生活小区等等。

目前已有许多涉及不同分布的总体分解方法，限于篇幅，本文仅讨论几种常见分布的总体分解方法，即正态及对数正态分解、指数分解、二项分解和泊松分解。首先介绍方法，然后展示一些地质实例，最后对总体分解的计算机程序作一简要说明。

一、混合总体的正态分解

这里讨论的虽然是混合总体的正态分解，但所有的结果均不难推广到对数正态的情形。

（一）混合总体的一维正态分解

设X为具有密度函数为$f(x)$的连续型随机变量，且$f(x)$被看成是k个正态子总体的混合，各子总体的均值、均方差和权系数分别为μ_1、σ_1、α_1、μ_2、σ_2、α_2、……、μ_k、σ_k、α_k，ϕ为正态密度函数，于是我们有混合总体一维正态分解的表达式：

$$f(x;\ \Theta)=\sum_{j=1}^{k}\alpha_j\phi\left(\frac{x-\mu_j}{\sigma_j}\right) \qquad (1)$$

$$\sum_{j=1}^{k}\alpha_j=1 \qquad (2)$$

式中

$$\Theta=(\mu_1、\sigma_1、\alpha_1、\mu_2、\sigma_2、\alpha_2……\mu_k、\sigma_k、\alpha_k)' \qquad (3)$$

在以后的一切总体分解模型中，如不特别申明，我们均假定式（2）是被满足的。

参数Θ可通过如下方法估计之。

1.图解法

辛克莱（1981）[1]是图解法的积极实践者。他用图解法正态或对数正态总体分解研究了大量矿床勘探与矿山开采的品位和储量数据，所识别出的子总体通常代表不同的矿化特征。

图解法总体分解是在概率纸上进行的。按所假设的分布不同，可使用正态概率纸或对

数正态概率纸；按总体叠加方式的不同，又可分为非包含混合概率曲线与包含混合概率曲线。于是，我们就有了四种基本的总体分解图解方法：

（1）非包含混合总体的正态分解；

（2）包含混合总体的正态分解；

（3）非包含混合总体的对数正态分解；

（4）包含混合总体的对数正态分解。

所谓非包含混合总体，指的是组成混合总体的子总体相互仅部分重叠或不重叠；包含混合总体则不然，它表现为组成观测总体的一个子总体完全被另一个子总体所包含。

图1至图4表示了图解法总体分解的四种基本模式。关于这四种模式的具体作图分解方法，可详见辛克莱的著作（1981）[1]。

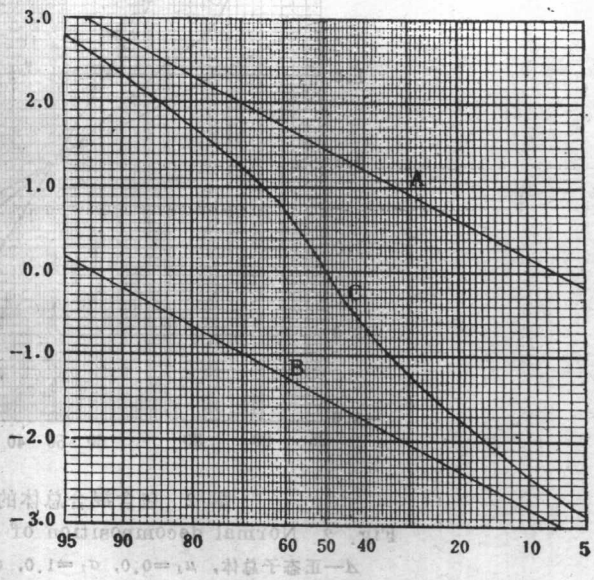

图1　非包含混合总体的正态分解

Fig. 1　Normal decomposition of a non-inclusion mixture.

A—正态子总体，$\mu_1=1.5$，$\sigma_1=1.0$，$\alpha_1=0.5$；

B—正态子总体，$\mu_2=-1.5$，$\sigma_2=1.0$，$\alpha_2=0.5$；

C—混合总体

图解法总体分解有许多优点，如直观、快速、简便易行，以及能处理截尾分布、三参数对数正态分布等。对于野外地质人员来说，不失为一种较好的方法。

2.矩量法

矩量法是进行总体分解的一种古老方法，其优点是原理简单，易于推导，无需使用复杂的数学工具；其缺点是矩量法的结果并不是个最优的解。1977年克拉克[14]对矩量法一维正态分解重新发生了兴趣，对它作了详细的叙述。

矩量法的实质在于建立观测总体的前若干阶矩同相应的子总体参数之间的关系。在一维正态总体分解中，首先计算观测总体的前若干阶k统计量，对估计观测矩是很方便的。

（1）计算k统计量

在这里，我们只打算用矩量法推导两个子总体的正态分解，因而将被估计的子总体参数为6，这样就需要计算前六阶观测矩和k统计量

据观测值可计算出

$$e_j=\sum_{i=1}^{N}(x_i-\bar{x})^j \quad j=2,3,6 \tag{4}$$

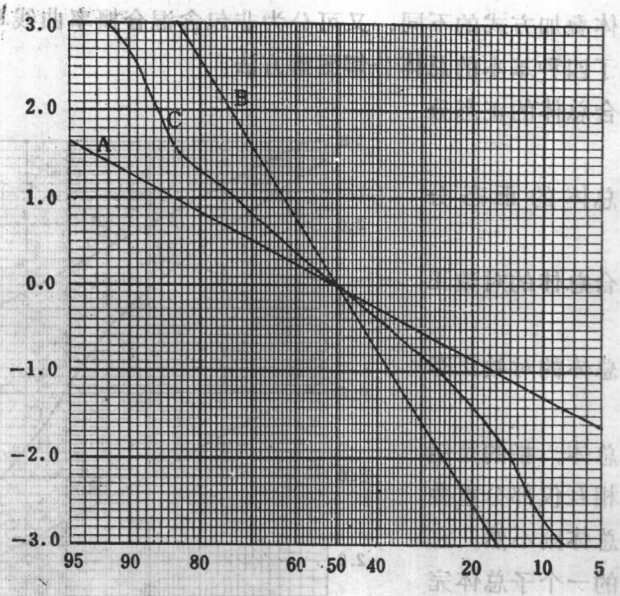

图 2 包含混合总体的正态分解

Fig. 2 Normal decomposition of an inclusion mixture.

A—正态子总体，$\mu_1=0.0$，$\sigma_1=1.0$，$\alpha_1=0.5$；

B—正态子总体，$\mu_2=0.0$，$\sigma_2=3.0$，$\alpha_2=0.5$；

C—混合总体

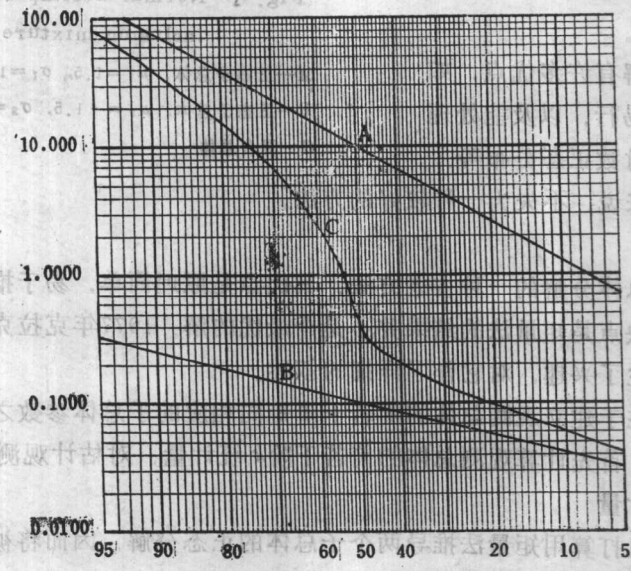

图 3 非包含混合总体的对数正态分解

Fig. 3 Lognormal decomposition of a non-inclusion mixture.

A—对数正态子总体，$\mu_1=10.0$，$\sigma_1=5.0$，$\alpha_1=0.5$；

B—对数正态子总体，$\mu_2=0.1$，$\sigma_2=0.5$，$\alpha_2=0.5$；

C—混合总体

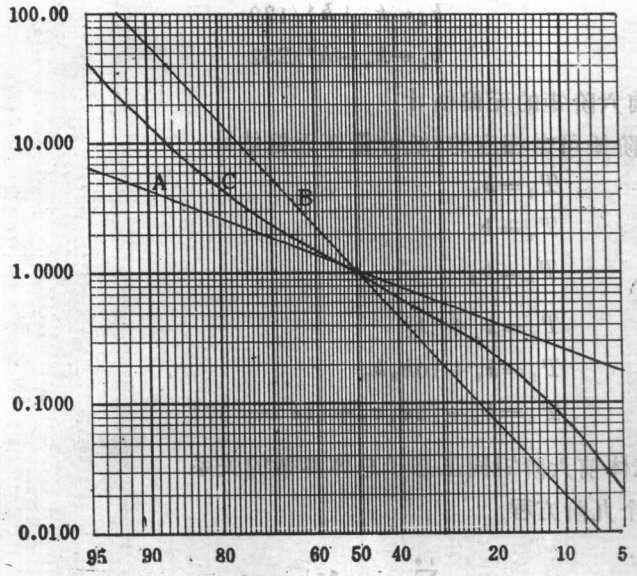

图 4　包含混合总体的对数正态分解

Fig. 4 Lognormal decomposition of an inclusion mixture.

A—对数正态子总体，$\mu_1=1.0$，$\sigma_1=3.0$，$\alpha_1=0.5$；

B—对数正态子总体，$\mu_2=1.0$，$\sigma_2=20.0$，$\alpha_2=0.5$；

C—混合总体

式中ϖ_i为第i个样品的观测值，$\overline{\varpi}$为子样均值，N为观测数。于是k统计量被定义为

$$
\begin{aligned}
k_1&=\overline{x}\\
k_2&=e_2/N^{(1)}\\
k_3&=e_3N/N^{(2)}\\
k_4&=[(N+1)e_4-3e_2^2(N-1)N^{-1}]N/N^{(3)}\\
k_5&=[(N+5)Ne_5-10(N-1)e_3e_2]N/N^{(4)}\\
k_6&=[(N^4+16N^3+11N^2-4N)e_6-15(N^3+2N^2-7N+4)\\
&\quad e_4e_2-10(N^3-2N^2+5N-4)e_3^2+30(N^2-3N+2)e_2^3]/N^{(5)}
\end{aligned}
\tag{5}
$$

式中

$$
N^{(j)}=\prod_{i=1}^{j}(N-i) \tag{6}
$$

如果上述k统计量是据每个原始样品计算的，则无须修正，但如果各阶矩是据归组后的直方图数据计算的，则需使用谢巴德法修正各偶数矩

$$
k_2\leftarrow k_2-h^2/12
$$

（7）

$$k_4 \leftarrow k_4 + h^4/120$$
$$k_6 \leftarrow k_6 - h^6/252$$

（2）观测总体前六阶矩的无偏估计

观测总体的前六阶矩可据相应的k统计量作无偏估计

$$\begin{aligned} T_1 &= k_1 \\ T_2 &= k_2 \\ T_3 &= k_3 \\ T_4 &= k_4 + 3k_2^2 \\ T_5 &= k_5 + 10k_3k_2 \\ T_6 &= k_6 + 15k_4k_2 + 10k_3^2 + 15k_2^3 \end{aligned} \qquad (8)$$

（3）建立观测总体前六阶矩同正态子总体参数的关系

首先定义一个九阶方程

$$\sum_{n=1}^{10} a_n p^{n-1} = 0 \qquad (9)$$

诸系数a取决于观测矩，由以下关系式确定

$$\begin{aligned} a_1 &= -24T_2^5 \\ a_2 &= 32T_3^4 b \\ a_3 &= 24T_3 c - 7T_3^2 b \\ a_4 &= 288T_3^4 - 12bcT_3 - b^3 \\ a_5 &= -2c^2 - 148T_3^2 c \\ a_6 &= 10b^2 - 24T_3 c \\ a_7 &= 36T_3^2 \\ a_8 &= -24b \\ a_9 &= 0 \\ a_{10} &= 24 \end{aligned} \qquad (10)$$

据皮尔逊[23]（1894），式中

$$\begin{aligned} b &= 9T_2^2 - 3T_4 \\ c &= 30T_2T_3 - 3T_5 \end{aligned} \qquad (11)$$

将（10）代入（9）解出p后，再解以下二次方程

$$g^2 - \gamma g + p = 0 \qquad (12)$$

式中系数γ由下式确定

$$s = (2T_3^2 - 2T_3bp - cp^2 - 8T_3p^2) / (4T_3^2 - bp - 2p^3) \qquad (13a)$$

$$\gamma=s/p \tag{13b}$$

设方程(12)的两个根分别为g_1及g_2，则两正态子总体的参数由下式算出

$$\begin{aligned}
\hat{\mu}_1&=T_1+g_1\\
\hat{\mu}_2&=T_1+g_2\\
\hat{\sigma}_1&=(T_2-T_3/3g_2-\gamma/3g_1+p)^{1/2}\\
\hat{\sigma}_2&=(T_2-T_3/3g_1-\gamma/3g_2+p)^{1/2}\\
\hat{\alpha}_1&=-g_2/(g_1-g_2)\\
\hat{\alpha}_2&=1-\hat{\alpha}_1
\end{aligned} \tag{14}$$

各子总体对观测总体的拟合程度有两种检验方法：其一是将由子总体组成的模型总体的前六阶矩同观测总体的前六阶矩进行对比；第二种是χ^2适度检验，它适用于具有直方图分组数据的检验。由于求解p时涉及一个九次方程，因此，除去虚根外，仍然会有多个解。通过对比前六阶矩或作χ^2检验，可从中选出一个最佳的解。两种检验方法所确定的最佳解往往是相同的。

3.非线性最小二乘法

德拉培和史密斯(N.R.Draper and H.Jr.Smith, 1967)[15]对非线性回归方法有非常详尽的叙述；巴德(Y.Bard, 1974)[3]发表了一个完善的非线性最小二乘法程序；麦克卡门(1969)[20]首次在地质学中介绍了这一方法；克拉克(1977)[7]则将它用于地质学中的一维总体正态分解问题。

(1)一元非线性回归

设Y与X为两个相关的随机变量，y_i与x_i分别为其观测值，则有回归模型

$$y_i=F(x_i;\ \Theta)+\varepsilon_i \tag{15}$$

式中

$$\Theta=(\theta_1,\ \theta_2\cdots\cdots,\ \theta_m)'$$

为回归模型的m个待定参数，ε_i为y_i的随机组分。如果$F(x;\ \Theta)$不能转换成关于Θ的线性函数，则上述模型就是一个非线性回归模型。

按最小二乘法准则，令

$$s=\sum_{i=1}^{N}(\varepsilon_i)^2=\sum_{i=1}^{N}[y_i-F(x;\ \Theta)]^2=\min \tag{16}$$

对于线性回归模型而言，通过对s关于诸θ_j的求导所得出的正好是个线性正规方程组，因此求解参数Θ当不成问题；但对于非线性回归模型而言，所得到的却是个非线性正规方程组，这时只有通过某种逐步线性化的方法来求解。在这里，我们使用高斯-牛顿法。

设Θ_0是对参数Θ的某种近似估计，则函数$F(x;\ \Theta)$在Θ_0点的泰勒展开式为

$$F(x;\ \Theta)=F(x;\ \Theta_0)+(\theta_1-\theta_{01})\frac{\partial F(x;\ \Theta_0)}{\partial\theta_1}$$
$$+(\theta_2-\theta_{02})\frac{\partial F(x;\ \Theta_0)}{\partial\theta_2}+\cdots+(\theta_m-\theta_{0m})\frac{\partial F(x;\ \Theta_0)}{\partial\theta_m}$$
$$+(\theta_j-\theta_{0j})\text{的高阶项} \tag{17}$$

如果Θ_0足够逼近Θ，则高阶项可忽略不计。于是关于Θ的非线性函数$F(x;\ \Theta)$被化为关于Θ的线性函数。这时再对目标函数s关于诸θ_j求导，于是获得线性正规方程组

$$D\cdot\Delta\Theta=g \tag{18}$$

式中g为常数向量，其元素为

$$g_j=\sum_{i=1}^{N}\frac{\partial F(x_i;\ \Theta_0)}{\partial\theta_j}[y_i-F(x_i;\ \Theta_0)] \tag{19}$$
$$j=1,\ 2,\ \cdots,\ m$$

D为系数矩阵，其元素为

$$d_{jl}=\sum_{i=1}\frac{\partial^2F(x_i;\ \Theta_0)}{\partial\theta_j\partial\theta_l}\qquad j,l=1,\ 2,\ \cdots,\ m \tag{20}$$

$\Delta\Theta$为解向量，其元素为

$$\Delta\theta_j=(\theta_j-\theta_{0j})$$

实际上，g涉及到非线性函数$F(x;\ \Theta)$在Θ_0点的一阶偏导数，而D则涉及到$F(x;\ \Theta)$在Θ_0点的二阶偏导数。因此，定义以下两个矩阵对形成正规方程组（18）的常数和系数项是很有利的。

A.雅可比矩阵　因为这里涉及的是一个一元非线性方程，故其雅可比矩阵是一个向量，由$F(x;\ \Theta)$关于Θ的一阶偏导数组成

$$DF(x;\ \Theta)=\begin{Vmatrix}\dfrac{\partial F(x;\ \Theta)}{\partial\theta_1}\\[2ex]\dfrac{\partial F(x;\ \Theta)}{\partial\theta_2}\\\vdots\\\dfrac{\partial F(x;\ \Theta)}{\partial\theta_m}\end{Vmatrix} \tag{21}$$

B.海赛矩阵　由$F(x;\ \Theta)$关于Θ的二阶偏导数组成

$$HF(x;\ \Theta)=\begin{Vmatrix}\dfrac{\partial^2F(x;\ \Theta)}{\partial^2\theta_1} & \dfrac{\partial^2F(x;\ \Theta)}{\partial\theta_1\partial\theta_2} & \cdots & \dfrac{\partial^2F(x;\ \Theta)}{\partial\theta_1\partial\theta_m}\\[2ex]\dfrac{\partial^2F(x;\ \Theta)}{\partial\theta_2\partial\theta_1} & \dfrac{\partial^2F(x;\ \Theta)}{\partial^2\theta_2} & \cdots & \dfrac{\partial^2F(x;\ \Theta)}{\partial\theta_2\partial\theta_m}\\\cdots & \cdots & \cdots & \cdots\\\dfrac{\partial^2F(x;\ \Theta)}{\partial\theta_m\partial\theta_1} & \dfrac{\partial^2F(x;\ \Theta)}{\partial\theta_m\partial\theta_2} & & \dfrac{\partial^2F(x;\ \Theta)}{\partial^2\theta_m}\end{Vmatrix} \tag{22}$$

有了雅可比矩阵和海赛矩阵在Θ_0点的值，就可以很快形成正规方程组（18）。

求解出参数增量$\Delta\Theta$后，通过下式

$$\Theta_1=\Theta_0+\Delta\Theta \tag{23}$$

求出Θ_1，把它作为对参数Θ的新的逼近。然后将Θ_1作为新的Θ_0，据此建立新的正规方程组以求取下一次逼近Θ的增量$\Delta\Theta$。重复上述过程，直到获得对Θ的足够精确的估计为止。

（2）一维总体正态分解的非线性回归模型

一维观测混合总体的分布函数为

$$Q(x;\ \Theta)=\sum_{j=1}^{k}\alpha_j\Phi\left(\frac{x-\mu_j}{\sigma_j}\right) \tag{24}$$

式中Φ为正态子总体累积分布函数，其参数由式（3）所定义。

由于在这里涉及到使用观测直方图数据进行总体分解，于是一维总体正态分解的非线性回归模型化为

$$\pi_i=F(x_i;\ \Theta)+\varepsilon_i \tag{25}$$

式中π_i为直方图上第i个区间的相对观测频率，$F(x_i;\ \Theta)$为关于Θ的非线性函数，其定义为

$$F(x_i;\ \Theta)=Q(x_i;\ \Theta)-Q(x_{i-1};\ \Theta) \tag{26}$$

求解Θ的最小二乘法准则为

$$S=\sum_{i=1}^{N}[\pi_i-F(x_i;\ \Theta)]^2=\min \tag{27}$$

$F(x_i;\ \Theta)$的一阶偏导数为

$$\frac{\partial F(x_i;\ \Theta)}{\partial\theta_j}=\frac{\partial Q(x_i;\ \Theta)}{\partial\theta_j}-\frac{\partial Q(x_{i-1};\ \Theta)}{\partial\theta_j} \tag{28}$$

$$i=1,\ 2,\ \cdots\cdots,\ N;\ j=1,\ 2,\ \cdots,\ 3k$$

$Q(x;\ \Theta)$对各参数的偏导数为

$$\frac{\partial Q(x;\ \Theta)}{\partial\alpha_j}=\Phi\left(\frac{x-\mu_j}{\sigma_j}\right)-\Phi\left(\frac{x-\mu_k}{\sigma_k}\right) \tag{29a}$$

$$j=1,\ 2,\ \cdots,\ k-1$$

$$\frac{\partial Q(x;\ \Theta)}{\partial\mu_j}=\frac{\alpha_j}{\sigma_j}\phi\left(\frac{x-\mu_j}{\sigma_j}\right)\quad j=1,\ 2,\ \cdots,\ k \tag{29b}$$

$$\frac{\partial Q(x;\ \Theta)}{\partial\sigma_j}=\frac{\alpha(x_j-\mu_j)}{\sigma_j^2}\phi\left(\frac{x-\mu_j}{\sigma_j}\right) \tag{29c}$$

$$j=1,\ 2,\ \cdots,\ k$$

利用（28）和（29），不难建立非线性函数（26）的雅可比矩阵和海赛矩阵，据此可形成正规方程组（18）。实施由（18）和（23）组成的迭代过程，可获得对参数Θ的最终估计。Θ即是式（3）定义的正态子总体参数。

（3）一维观测总体的对数正态分解

如果有理由认为观测总体是若干个对数正态子总体的混合，则只需将原始观测值转换为$\ln x$，然后对$\ln x$实行上述总体正态分解步骤就行了。但这时所估计出的参数μ_j与σ_j^2是子总体的对数均值与对数方差。其"真实"的均值λ_j与方差w_j^2由下式确定

$$\lambda_j=\exp\left(\mu_j+\tfrac{1}{2}\sigma_j^2\right) \tag{30a}$$

$$\omega_j^2=\lambda_j^2[\exp(\sigma_j^2)-1] \tag{30b}$$

$$j=1,\ 2,\ \cdots,\ k$$

在完成了某种模型的总体正态分解后，可用χ^2检验来考察模型分布对观测直方图的拟合度。

（二）混合总体的多维正态分解

德依（Day，1969）[14]把正态总体分解推广到多维情形。这时式（1）的总体分解模型的多变量比拟为

$$f(x;\ \Theta)=\sum_{j=1}^{k}\alpha_j\phi[x-\mu_j]'\Sigma^{-1}(x-u_j)] \tag{31}$$

式中x为p维随机向量X的观测值，μ_j为第j个子总体的均值向量，Σ^{-1}为公共的协方差矩阵的逆。多维总体分解的目的在于估计参数

$$\Theta=(M',\ \Sigma,\ A) \tag{32}$$

其中M为$p\times k$的均值矩阵

$$M=\begin{pmatrix}\mu_{11}\ \mu_{12}\cdots\mu_{1k}\\ \mu_{21}\ \mu_{22}\cdots\mu_{2k}\\ \cdots\cdots\cdots\\ \mu_{p1}\ \mu_{p2}\cdots\mu_{pk}\end{pmatrix} \tag{33a}$$

Σ为$k\times k$的公共协方差矩阵

$$\Sigma=\begin{pmatrix}\sigma_{11}^2\ \sigma_{12}^2\cdots\sigma_{1k}^2\\ \sigma_{21}^2\ \sigma_{22}^2\cdots\sigma_{2k}^2\\ \cdots\cdots\cdots\\ \sigma_{k1}^2\ \sigma_{k2}^2\cdots\sigma_{kk}^2\end{pmatrix} \tag{33b}$$

A为$p\times 1$的子总体权系数向量

$$A'=(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_k) \tag{33c}$$

上述参数可通过最大似然估计获得。据N个观测值产生的混合正态分布的最大似然函数为

$$L(M,\ \Sigma,\ A)=(2\pi)^{-\frac{Np}{2}}|\Sigma|^{-N}\prod_{i=1}^{N}\left\{\sum_{j=1}^{k}\alpha_j\exp\left[-\frac{1}{2}(x_i-\mu_j)'\Sigma^{-1}(x_i-\mu_j)\right]\right\} \tag{34b}$$

按最大似然法的例行步骤，将函数L取对数，然后对各参数取偏导数并令其为零，解由此而得到的线性方程组，即可得出对M、Σ、A的估计式，该估计式具有依赖于后验概率的形式

$$\hat{\alpha}_j=\frac{1}{N}\sum_{i=1}^{N}\hat{P}(j|x_i) \qquad j=1,\ 2,\cdots,k \tag{35a}$$

$$\hat{\mu}_j=[\sum_{i=1}^{N}x_i\hat{P}(j|x_i)]/[\sum_{i=1}^{N}\hat{P}(j|x_i)] \tag{35b}$$

$$\hat{\Sigma}=\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{k}(x_i-\hat{\mu}_i)(x_i-\hat{\mu}_j)'\hat{P}(j|x_i) \tag{35c}$$

而后验概率又具有依赖于判别得分的形式

$$\hat{P}(j|x_i)=\exp[\hat{d}_j(x_i)]/\left\{\sum_{h=1}^{k}\exp[\hat{d}_h(x_i)]\right\} \quad j=1,2,\cdots,k \tag{36}$$

式中$\hat{d}_j$为据第j个判别方程估计的判别得分。

众所周知，判别得分由判别方程

$$\hat{d}_j(x_i)=\hat{c}_{j0}+\hat{c}_j'x_i \qquad \begin{matrix} j=1,2,\cdots,k \\ i=1,2,\cdots,N \end{matrix} \tag{37}$$

估计之。式中c_{j0}为第j个判别方程的常数项，c_j为第j个判别方程的判别系数向量。

最后，判别常数与判别系数又具有依赖于正态子总体参数Θ的形式

$$c_{j0}=-\tfrac{1}{2}\mu_j'\Sigma^{-1}\mu_j+\ln\alpha_j \qquad j=1,2,\cdots,k \tag{38a}$$

$$c_j'=\Sigma^{-1}\mu_j \qquad j=1,2,\cdots,k \tag{38b}$$

显然，关系式（35）、（36）、（37）、（38）组成了一个迭代过程。迭代一般由给定正态子总体参数初值开始，于是可据（38）获得对判别常数与系数的估计，下一步则据（37）估计判别得分，然后再据（36）估计后验概率，最后又据（35）估计出子总体的新参数，于是进入新的一轮迭代。如此重复进行，直到满足一定的迭代精度为止。

为了考查子总体的显著性，可计算各子总体间的马氏距离，并进行相应的F检验。

（三）混合总体的约束正态分解

正态总体分解要解决的实际上是一个正态分类最优化问题。将一组观测进行一维或多维正态分类，在大多数情况下是多解的，我们总是力图寻找出一个在地质上合理的最优正态分类决策。给正态总体分解过程加上某种约束条件，将有可能迫使总体分解朝最优的正态分类方向进行。

在实施多元正态分解时，如果令分解的结果满足于某个重要变量的单变量正态分解结果，换言之，令多元正态分类在某个单变量正态分解的子总体直方图的约束下进行，则将会明显地抑制其多解性，并通常能得到一个令人满意的最优解[18]。

在二元正态总体分解中，这种单变量直方图约束对解的改善尤为明显，特别是在进行单变量的空间图形识别时，混合总体的约束二元正态总体分解将提供一个最优化的空间图形识别工具。

设有一组空间分布的单变量观测数据，要通过约束二元正态总体分解来识别隐含在这些数据中的空间图形，其具体步骤如下：

1.建立单变量正态子总体直方图

单变量观测直方图的相对频数为

$$w^*(c,t)=w(c,t)/N \qquad c=1,2,\cdots,t \tag{39}$$

式中c为直方图上的区间号，t为直方图区间总数，$w(c,t)$为由t个区间组成的直方图上落入第c个区间的观测数（即绝对频数），$w^*(c,t)$为相对频数，N为观测总数。

通过一维正态总体分解，可求得各正态子总体参数，于是观测的相对频率可近似地表达为

$$\pi^*(c,t)=\sum_{j=1}^{k}\int_{x_{c-1}}^{x_c}a_j\phi\left(\frac{x-\mu_j}{\sigma_j}\right)dx+\varepsilon(c,t) \qquad (40)$$

$$c=1,\ 2,\ \cdots,\ t$$

式中x_c与x_{c-1}为直方图第c个区间的上、下限，ε为拟合误差。于是第c个区间中第j个子总体的理论绝对频数为

$$n_j(c,t)=\frac{\int_{x_{c-1}}^{x_c}a_j\phi\left(\frac{x-\mu_j}{\sigma_j}\right)dx}{\sum_{h=1}^{k}\int_{x_{c-1}}^{x_c}a_h\phi\left(\frac{x-\mu_h}{\sigma_h}\right)dx}\cdot\eta(c,t) \qquad (41)$$

$$c=1,2,\cdots,t$$

由于存在拟合误差$\varepsilon(c,t)$，故第c个区间所有k个子总体的理论绝对频数之和

$$\sum_{j=1}^{k}n_j(c,t)$$

通常并不正好等于相应的观测绝对频数$\pi(c,t)$。这时我们需要把第c个区间的拟合误差$\varepsilon(c,t)$平差到每个$n_j(c,t)$中，于是得到修正后的第c个区间第j个子总体的绝对频率为

$$n'_j(c,t)=n_j(c,t)+\varepsilon_j(c,t) \qquad c=1,\ 2,\ \cdots,\ t \qquad (42)$$

并要求$n'_j(c,t)$为正整数．且

$$\varepsilon(c,t)=\sum_{j=1}^{k}\varepsilon_j(c,t) \qquad c=1,\ 2,\ \cdots,\ t \qquad (43)$$

式中$\varepsilon_j(c,t)$为第c个区间第j个子总体的拟合误差。这样一来，$n'_j(c,t)$就完全拟合$\pi(c,t)$了，即

$$\sum_{j=1}^{k}n'_j(c,t)=\pi(c,t) \qquad c=1,\ 2,\ \cdots,\ t \qquad (44)$$

2.构造邻域背景变量

通过单变量正态总体分解，可以将每个观测归入到某个正态子总体中，即对每个观测进行正态归类，然后将观测按所属的子总体不同以不同的符号标在图上，就形成了一张类型分布图。由于我们只考虑了原始变量本身，这种单变量总体分解所识别出的空间图形往往连续性不好、噪声干扰大。为了得到一张更为连续的空间图形，可以用某些已有的圆滑方法。本文使用了斯威扎（Swaizer，1981，1982）[24,25]的办法。他在应用多变量统计方法识别空间图形时，不仅考虑了原始变量本身，还考虑了由原始变量派生出来的扩增变量——原始变量的邻域平均值。将斯威扎的方法用于解决我们的问题，于是单变量图形识

别就化为一个双变量图形识别问题，第二个变量是第一个变量的邻域平均值。

设原始观测变量为X_1，任一邻域内原始观测变量的平均值为X_2，则一个二元随机向量被定义为

$$X=\begin{Bmatrix}X_1\\X_2\end{Bmatrix} \tag{45}$$

邻域平均值X_2被定义为

$$X_2=E[X_1(A)] \qquad X_1(A)\in X_1 \tag{46}$$

式中$X_1(A)$是被定义在地理域A内那些原始变量观测值的集合，是X_1的子集，A是观测点位置(u, v)的函数，即

$$A=F(u, v) \tag{47}$$

对于任意位置(u, v)，原始变量观测值$x_1(u, v)$位于A的中心，但$X_1(A)$不包含$x_1(u, v)$。

对规则网格观测而言，A是由单元(u, v)为中心的周围8个、24个、48个或更多个单元所构成的范围。以8单元邻域为例，单元(u, v)的邻域平均值为

$$\begin{aligned}x_2(u,v)=[&(x_1(u-1,v-1)+x_1(u-1,v)+x_1(u-1,v+1)\\&+x_1(u,v-1)+x_1(u,v+1)+x_1(u+1,v-1)\\&+x_1(u+1,v)+x_1(u+1,v+1)]/8\end{aligned} \tag{48}$$

对不规则观测，可按如下方式定义地理域A：设A是一个以点(u, v)为圆心，半径为r的圆，p是任意一具有位置(u', v')的点，p点到圆心的距离为d_p，则A是所有满足

$$0<d_p\leqslant r \tag{49}$$

的点的集合。在实际问题中，我们总是只能获得有限个观测，因此$X_1(A)$可根据所有满足于(49)的观测来计算之。

3. 建立分类矩阵

样品分类矩阵被定义为

$$\Delta=\begin{Bmatrix}\delta_{11} & \delta_{12} & \cdots & \delta_{1N}\\ \delta_{21} & \delta_{22} & \cdots & \delta_{2N}\\ \cdots & \cdots & \cdots & \cdots\\ \delta_{k1} & \delta_{k2} & \cdots & \delta_{kN}\end{Bmatrix} \tag{50}$$

式中

$$\delta_{ji}=\begin{cases}1 & \text{当观测属于第}j\text{个子总体时}\\ 0 & \text{其他情形}\end{cases}$$

双变量约束正态总体分解使用一套迭代算法，迭代由分类矩阵赋初值开始。分类矩阵的初值最好在某种先验知识的指导下进行。