

数学地质专辑

(一)

刘 承 祚 等著

509

地 质 出 版 社

数学地质专辑

(一)

刘承祚 等著

地质出版社

内 容 提 要

本册收集了十九篇文章,包括数学地质的数学基础、数学地质方法和在地质学各分支的应用两大类。“数学地质现状与任务”一文则比较全面地将数学地质做了介绍,使读者对这一学科有个完整的概念。

“数学地质专辑”刊载国内外数学地质的学术论文和工作经验,以后将陆续分册出版。

本册供数学地质工作者、金属和非金属找矿地质人员、煤田地质人员、石油地质人员、物化探人员、矿山地质工作者,以及相应各院校师生参考。

2166/59

数 学 地 质 专 辑

(一)

刘承祚 等著

*

地质部书刊编辑室编辑

地质出版社出版

(北京西四)

地质印刷厂印刷

(北京安德路47号)

新华书店北京发行所发行·各地新华书店经售

*

开本 $787 \times 1092^{1/16}$ ·印张13·插页2·字数319,000

1980年5月北京第一版·1980年5月北京第一次印刷

印数1—4,975册·定价1.70元

统一书号: 15038·新501

目 录

数学地质的现状和任务	刘承祚等 (1)
聚类分析的方法及其应用	方开泰 (13)
关于数量化理论 I、II 的数学模型	周光亚等 (24)
地质统计分析方法简介及其应用	於崇文等 (31)
有序地质量最优分割问题	姜元仁等 (44)
一个前寒武纪变质铁矿东矿带地质特征及大比例尺统计预测初步研究	赵鹏大等 (53)
海南石碌铁矿两类变质岩石的因子分析及地质意义	潘恩沛等 (65)
应用多元统计分析方法对封三洞角砾岩筒型斑岩铜矿的研究	孙树浩 (76)
云南某拗陷含盐盆地及钾盐矿床的多元统计分析	云南大学数学系 云南省地质局物探队 (90)
某些统计方法在煤田地质中的应用	门桂珍 (102)
数学方法处理红外光谱资料在评价生油岩中的应用	岑梅等 (114)
华南燕山期含矿花岗岩成分特征的多元统计研究	孙文鹏等 (124)
数学地质在研究岩浆岩及其有关铁矿方面的初步应用	秦方鸿 (138)
利用因子分析推断陕甘宁盆地南部延长统沉积时期的蚀源区	唐文松等 (145)
信息函数与古生态研究	汪品先等 (154)
金云母折光率的找(铁)矿意义	吴思本 (163)
趋势面分析用于断裂构造研究一例	唐茂绍 (172)
利用电子计算机处理水化资料实例	李文兴等 (178)
地质数据库探讨	于志钧 (191)
一九七八年十月全国数学地质学术讨论会论文题目目录	(200)
编者的话	(205)

数学地质的现状和任务

刘承祚 孙惠文

地质学是一门有悠久历史的自然科学基础学科。长久以来,地质工作者主要靠记录和描述的方法收集实际地质资料,通过分析和归纳得出地质学的主要规律和理论。地质学长期以来基本上是一门定性的科学,在地质学研究中很少采用定量方法,并且很少应用数学。伟大革命导师马克思、恩格斯都曾指出过数学在自然科学研究中的重要意义。马克思指出:“只有科学应用了数学才达到完善程度”(见拉法格著:“回忆马克思恩格斯”,第七页,人民出版社,1973年),恩格斯在“反杜林论”一书中曾指出:“要确立辩证的同时又是唯物主义的自然观,需要具备数学和自然科学的知识。”毛主席也曾指出:“胸中有‘数’。就是说,对情况和问题一定要注意到它们的数量方面,要有基本的数量的分析。”在地质学中应用数学大约已有150年历史,但最初只有个别的事例,只有在地质学中推广应用电子计算机以后,才开始比较普遍地和大量地应用数学,并逐渐形成了一个新学科——数学地质。

一、数学地质的定义和主要研究内容

数学地质是在地质和数学互相渗透、紧密结合的基础上产生的一门边缘学科,它是应用数学方法研究地质学基础理论和解决地质学中实际问题的地质学分支,电子计算技术是数学地质研究的主要技术手段。

数学和地质的结合是指应用数学各个分支领域的成果,解决地质学的理论问题和实践问题。数学是一门研究数与形的科学,数学中研究数的部分属于代数学的范畴,研究形的部分属于几何学的范畴。数学有着极其丰富的内容。目前在地质学中只应用了数学的很小一部分。应用较多的是随机类学科(概率论、数理统计、随机过程论以及多元统计),函数论、微分方程、数学分析等,在地质过程的数学模拟中应用较多,拓扑学在地质学中应用也不少。总之,数学在地质学中尚有非常广阔的应用领域。

在现阶段,数学地质的主要内容包括地质多元统计,地质过程的数学模拟,地质数据的储存、索取、自动处理和显示,地质统计学。

1. 地质多元统计

应用各种多元统计方法研究地质问题是数学地质迄今为止的一个主要方面。已经应用的多元统计方法有:回归分析、趋势面分析、判别分析、聚类分析、主成分分析、因子分析、对应分析、典型相关和典型回归、非线性映象分析等,而且不断地向地质学中引进新的多元统计方法。由于地质问题所具有的多元特性,多元统计方法几乎已应用于地质学的所有分支,而且都取得了一定效果。

2. 地质过程的数学模拟

应用数学方法模拟地质过程已是研究基础地质理论的重要途径之一,是数学地质的另一个重要内容。通过数学模拟在电子计算机上使地质过程再现,充分利用电子计算机内存大、速度快的特点,提高地质过程模拟研究的效率,缩短地质过程模拟研究的周期,会大大提高地质理论研究的进程,甚至所得成果,有可能从根本上改变某些旧的地质结论。地质过程的数学模拟研究可分为两大类:第一类为确定性数学模拟,即应用精确的可以得到确定解的数学方法模拟地质过程;第二类为随机过程模拟,应用随机过程方法模拟地质过程。近年来又出现了将确定性模拟和随机模拟结合起来的模拟方法,是数学模拟的新进展。

3. 地质数据的储存、索取、自动处理和显示

应用电子计算机储存、索取、自动处理和显示地质数据是一个很重要的研究方面。这方面的研究成果,可以使地质资料高度集中,避免资料管理的重复,高速、可靠地按工作需要提供资料,使地质工作从数据到成图过程自动化,并相应地要求改变某些传统地质工作方法,要求地质记录标准化,地质工作定量化,样品分布合理化,测试技术精密化等。

4. 地质统计学

地质统计学是由南非矿山工程师克里金(D. G. Krige)首先创立的,由法国马赛隆(G. Mathron)加以完善和发展的一门数学地质独立分支学科,主要涉及到矿产储量计算、品位计算及计算误差问题。近年来在马赛隆及其领导的小组积极工作下,地质统计学工作有显著的进展。

随着数学向地质学进一步渗透,数学地质这门新学科所包括的内容将会越来越广泛。近来,将微分拓扑学的研究成果应用于地质学而建立了地质学中的突变理论模型就是其例。突变理论是数学中的一个最新进展,是法国著名数学家勒内·汤姆于1968年提出的一种新观点。这种理论试图用数学工具描写那种灾难性的或突如其来的变化现象。这种数学上的最新进展很快应用于研究地质学基本理论问题,并取得了成效。如应用突变理论研究断层运动,把断层运动看作一个突变过程。又如应用突变理论研究二叠纪海洋无脊椎动物的灭绝,提出大洋中盐分减少是造成二叠纪海洋无脊椎动物灭绝的重要因素。

地质学的数学化问题也是数学地质的基本内容之一,研究如何用数学的语言和公式表达地质学的定义、概念和基本规律,从而将地质学建立在数学化的基础上。

二、数学地质发展简史

地质学和数学开始结合到数学地质产生的漫长过程可以分为以下五个阶段:

第一阶段 由1840年到1935年。这一时期的主要特点是在地质学中应用数学的初步尝试,和在个别方面中少量的分散研究。1840年,莱伊尔(Lyell)通过古生物化石的统计分析对第三系地层进行了划分。1890年,卡尔·皮尔逊(Karl Pearson)编写了“数学进化论贡献”一套丛书,内有古生物化石的统计分析。俄国人列文森-列星格(Левинсон-Лессинг)在1914年至1934年期间,通过研究岩石的岩浆系数的频率分布,研究了安山岩、玄武岩、流纹岩的分类。1929年,勃林克曼(R. Brinkmann)作了一些生物地层学方面的统计研究工作。

第二阶段 由1936年到1945年。这一阶段的特点是数学方法的应用由地质学的个别问

题逐渐扩展至地质学的一些分支。1939年,西姆波森(G. G. Simpson)等编著了“定量动物学”一书,为后来古生物统计学的发展奠定了基础。美国人克鲁拜因(W. C. Krumbein)自1934年开始进行沉积作用和地层的统计分析,成为美国数学地质的奠基人和一个学派的领导者。1944年,苏联维斯捷列乌斯(А. Б. Вистелиус)在苏联科学院报告集上发表了“分析地质学”一文,提出用定量方法研究地质问题的初步思想。从此他本人从事了三十多年的数学地质工作,成为苏联数学地质的创始人和国际数学地质学会的第一任主席。在这一阶段中,电子计算机刚刚研制成功,进行了少量试制,第一台电子计算机Z3 1941年在德国诞生。这时,电子计算机在地质学中的应用还没有提到日程上来。

第三阶段 自1945年到1960年。这一阶段的主要特点是数学方法普遍应用于地质学所有分支和单变量、双变量统计方法的普遍应用。在这一阶段中,电子计算机开始引用到地质学中来。1940年,苏联有人研究金属矿床元素统计分布特点。1954年,绍(D. M. Shaw)等人应用统计方法研究地球化学问题;1956年,切叶思(F. Chayes)应用均值、方差、标准差于岩石学研究;1958年,克鲁拜因从事区域地质统计分析方面的工作。美国宾夕法尼亚大学于1946年建立了ENIAC电子计算机,1952年建立了数学绘图仪,1953年编成了第一个FORTRAN编译器,1954年首次成批生产IBM-650电子计算机,1958年,克鲁拜因首次在地质杂志上公布电子计算机地质计算程序,地质学开始了应用电子计算机。1958年开始生产第二代晶体管电子计算机,同时在一些国家中开始应用ALGOL语言。

第四阶段 自1961年至1970年。在这期间,数学方法和电子计算机在地质学中开始广泛应用。美国亚利桑那大学自1961年开始召开了一系列“电子计算机在矿产工业中的应用”讨论会。1962年由于晶体管电子计算机的成批生产和广泛应用而导致数学地质文献数目激增。1963年宣布第三代微型电子计算机试制成功,美国堪萨斯地质调查所电子计算机程序集连续出版。在这一年,电子计算机在地质学中应用方面的文章第一次超过100篇。1964年,在美国达特蒙大学第一次成功地应用了电子计算机分时系统。堪萨斯地质调查所于1966年召开第一次“电子计算机在地球科学中的应用”讨论会(以后连续开了八次)。“美国石油地质工作者公报”杂志设立了“电子计算机应用”的专门编委。1967年,在美国石油地质工作者协会中建立了电子计算机数据储存和索取委员会,成立了国际地质科学联合会的地质数据储存、自动处理和索取委员会(COGEO DATA)。1968年在巴黎召开的国际地质会议上成立了国际数学地质协会(IAMG),开始出版国际数学地质协会杂志和地质计算程序公报,美国地质调查所首次公布其电子计算机贡献文集,“电子计算机在地球科学中应用”方面的第一本书出版。在这一阶段,多元统计方法在地质学中大量应用,数学地质发展成为独立的学科。

第五阶段 1971年至现在。这一阶段的特点是数学地质学科向更高的水平发展,地质过程数学模拟,地质数据储存、索取、自动处理和显示在数学地质中占据愈来愈重要的地位,愈来愈多的数学方法应用于地质学中,地质统计学取得明显进展,由法语国家向英语国家逐渐推广,并且水平不断提高,地质多元统计有形成独立分支的趋势。数学和地质学的不断结合推动了数学地质的进展。

在我国早期的地质工作中有应用数学的个别事例,但数学地质做为—门学科来对待在我国开始于六十年代初期。最初主要是应用因子分析、趋势面分析研究地层数据,1970年以后逐渐扩展至地质学其他分支。目前在我国相当多的研究机关、高等院校和生产单位中

都有人从事数学地质工作，一支数学地质工作队伍已初步形成。

三、数学地质工作的现状

数学地质是一个年轻的地质学科。与矿物学、岩石学、地层学、构造地质学、矿床学等学科相比，数学地质研究的对象和内容，还都不十分明确，有待逐步充实；研究使用的方法和技术手段，还都不十分成熟，有待逐步完善。数学地质是生命力很强的新学科。

在前一个时期，数学地质在国内外都以随机类数学在地质工作中的应用研究为主，地质问题的数学模拟虽是这一学科的一个重要研究方面，但从已发表的文献看，远远不及从事统计方面的人多。地质数据库从一开始就是数学地质中的一个独立分支，近年来发展很快。地质统计学已由少数几个国家逐渐推广到更多的国家中使用。

已经大量应用的数学地质方法正朝着提高的方向发展，这一方面表现在方法的改进，另一方面表现为个别理论问题研究的深入。以聚类分析为例，为了适合处理大量地质数据，在原有系统聚类分析的基础上，出现了动态聚类分析方法，同时深入研究了如何提高两个聚类之间分离程度的理论问题。

当前数学地质发展的一个重要特点是向地质学中引进数学的速度很快，一些数学上的重大突破，在数学地质中很快便有所反映，引进的速度之快，有时是令人吃惊的。

地质计算程序的系统报导，仍是一个推进数学地质工作的重要方法。对数学地质的发展，美国堪萨斯地质调查所在1966—1971年期间发表的地质计算程序（共50期）曾起了重要作用。目前国际上有两个数学地质杂志，一是“国际数学地质学会志”，专门发表数学地质的研究成果；另一个是“电子计算机与地球科学”，是发表地质计算程序的快报。

在个别的地质学分支中，已出现用数学从根本上改造地质学的苗头。如在地层学研究中出现了“形式化地层学”（formalized stratigraphy），其内容主要是应用严格的数学术语来表达地层学的原理和概念。为了便于用电子计算机处理地层数据，使用一组固定的符号来表达地层学的记录，这便为地层学进一步利用数学和电子计算机开辟了广阔的前景。

下面我们将数学方法和电子计算机在地质学各个分支中的应用现状作一概括介绍。

1. 在构造地质学中的应用

数学方法和电子计算机在构造地质学中的应用大致可分为以下几个方面：

（1）构造要素和构造形态的定量化描述和制图

褶皱轴的方位一般在野外进行测定，或用史密特等面积投影的下半球中的 β -或 π -图来测定，对后者已经编有很多专门计算程序，可以在电子计算机上实现。

在构造形态的研究方面，应用电子计算机自动从构造等值线中勾出地层的起伏，应用各种数字滤波器由地层的总的构造变异性中分出局部构造因素，并进一步分析这些因素的实质。

趋势面分析是研究构造形态的重要方法，即应用了多项式趋势面，又应用了富氏序列趋势面。趋势面分析可以把构造地质数据分解成两部分，一部分为大区域的构造形态，另一部分为局部的构造形态，然后再对这两种构造形态分别进行深入研究。

利用电子计算机可以自动绘出构造要素和构造形态图件。在绘制构造等值线图的程序

中, 普遍采用了 Spline 函数拟合方法, 此法比多项式函数拟合能划出更准确的构造等值线图。不少作者编了计算 Spline 函数和绘制等值线图的程序可供采用。

(2) 构造形态及其影响因素的多元统计研究

构造形态及影响构造形态的各种地质因素之间的关系可以用多元统计方法进行研究。例如可以用回归分析方法研究褶皱形态与岩性、岩层厚度、区域变质程度之间的关系。

(3) 构造地质数据的储存和索取

为了构造地质研究可以设计专门的数据文件和数据库。如为了研究变质岩地区构造地质问题而设计的专门数据文件等。

(4) 构造地质问题的数学模拟

首先应当指出有限单元法在构造地质模拟中的应用, 此法已用于模拟岩体中应力分布, 进而评价岩体的工程地质性质, 计算区域构造应力场, 研究局部构造和大区域构造。有人应用数学物理方程研究大洋中脊岩流上涌运动, 并进一步探讨板块运动的驱动力问题; 应用随机线、随机多边形以及球面上的随机大圆模拟断裂构造。

总之, 目前数学方法和电子计算机在构造地质中的应用还不多, 有广阔的发展前途。

2. 在地层学和沉积学中的应用

将拓扑学应用于地层学研究, 一般是将实际观测的地层系统和理论模型系统用拓扑学方法联系起来, 将观测到的地层数据转换成地质环境和古地理的概念。在这类研究中, 理论系统是由地质环境的地层模型和关于地质环境的理论概念组成的, 实际观测系统是由地层岩性的参数总体组成的。由拓扑学和地质学定义给出参数模型, 应用一定方法对模型进行研究可以求出一个地区的古地理面貌。

多元统计方法广泛应用于地层的划分和对比。马氏链用于模拟地层层序的形成。

在沉积学研究中数学方法和电子计算机也得到了广泛应用。例如应用主成分分析、聚类分析等方法对沉积相进行多元统计分析, 取得了一定效果。应用蒙特卡洛模拟方法对罗马尼亚东部喀尔巴阡山复理式沉积进行了模拟, 阐明了这种地层的沉积机制。应用条件模拟方法对巴黎盆地的二叠纪和三叠纪沉积旋回进行了三维数学模拟, 为阐明这一沉积旋回的成因提供了依据。应用电子计算机模拟模型研究了风力、风向变化对于不同类型砂丘形成和迁移的影响, 同时详细研究了沙丘侵蚀和堆积的机制。数学地质工作者应用典型相关方法研究了非洲西部三角洲沉积物中 Pb 和 Zn, Eh 值和有机成分, 碳酸盐贝壳以及 Mn、P、S 含量之间的相关关系。

对沉积盆地地下水流动的数学模拟, 除了能帮助评价水资源以外, 也能为沉积盆地本身的结构提供很多有益看法。

近年来在沉积学量化研究中采用了半马尔可夫过程。通常把沉积过程分解为两个阶段, 第一个阶段是由地质环境控制总的岩性特征, 第二个阶段为形成相应的层理。前一个过程用马氏链模拟, 后一个过程用半马尔可夫过程模拟, 取得了较好效果。这是沉积作用数学模拟取得的最新进展。

3. 在古生物学中的应用

在古生物学中的应用主要表现在以下几个方面。

统计方法在古生物学研究中, 主要应用了方差分析、判别分析、因子分析以及其他多元统计方法。

古生物化石和生物相的数字分类,对按地层顺序排列的化石数据可以用时间序列分析进行研究。

在化石形态模拟方面可将生物的形态看作一个动态的生长过程的结果,因此动态模型最适于进行化石形态的模拟。应用模拟计算机可以模拟腹足纲化石贝壳的整个形态。在模拟过程中可用改变一些参数的方法,改变贝壳的形态。适当地改变并重新组合以上几个参数,可以得出贝壳的不同形状和大小。在数字计算机上也可以作化石形态的模拟,但要花费较多的计算时间。

应用电子计算机做古生物化石的座标转换,比用手工做座标转换效率高,又可避免主观性。

电子计算机模拟、座标转换和机器绘图构成了电子计算机古生物学的重要内容。

4. 在地球化学研究和化探工作中的应用

多元统计方法在地球化学研究中得到了广泛应用。从在地球化学中应用的角度,可以将多元统计方法分为两大类。第一类主要是属于数字分类的一些方法(如判别分析、Q型因子分析、聚类分析)。这些方法主要用于研究标本之间的关系,有的同时也能用于研究变量之间的关系。这一类方法已比较成功地用于解决各种地球化学问题。第二类是用于研究变量之间关系的多元统计方法(相关分析、回归分析、R型因子分析、典型相关分析)。这一类方法可用于研究化学元素的共生组合关系。但这些方法在地球化学研究中的应用遇到了一些困难。由于地球化学工作者常常利用“组成变量”,这些变量具有“常数和”,或称为“闭合数据”。“常数和”是指例如由伟晶岩脉中取出的只包括石英和长石两种矿物的样品,则这两种变量用比例或百分数表示,显示出一种完全的负相关。由于对所有样品这两个变量的和都是常数,所以一个变量百分数的增长必然导致另一个变量百分数的减少,因此必然是负相关。但这种负相关仅仅是数据“常数和”作用的结果。目前国际上对数据的“常数和”问题作了一定研究,但尚未彻底解决。

数学方法和电子计算机已应用地球化学探矿工作的全过程:由化探工作的设计(包括合理取样网度选择),化探样品的物质成分分析,一直到成果的解释,都利用数学方法在电子计算机上进行。

应用概率统计方法研究化探取样的合理网度开始于五十年代末期,但最初提出的方法由于需要很多变量和参数,不便于实际应用。近年来提出了一些新的计算合理取样网度的概率统计方法,这些方法所需参数较少,且已作成诺模图,查检迅速,使用方便,便于实际应用。

数学方法和电子计算机在化探工作中的应用还表现在以下方面:应用数理统计方法处理样品中金属元素含量的快速分析结果,应用数学方法评价化探数据分析误差和测量误差,应用数理统计方法编制铜矿带和多金属成矿带的地球化学图和成矿预测图,应用趋势分析方法编制地球化学图和地球化学剖面,应用数理统计方法解释地球化学次生晕,对地化异常进行分类,建立矿体及其晕圈分带性的数学模型,用主成分分析法解释化探数据,应用数学方法对物探和化探数据进行综合分析解释,化探数据数学处理的理论基础研究等。

5. 在岩石学和矿物学中的应用

主要有两个发展方向。

第一个发展方向是应用数理统计方法,特别是多元统计方法研究岩浆岩、变质岩以及

各种经过交代改造的岩石,对各类岩石进行数字分类,特别是对难于用传统方法分类的岩石进行数字分类。对岩浆岩的岩石化学进行多元统计研究。研究各类矿物的共生组合规律,阐明矿物学的一些理论规律。

第二个发展方向是成岩作用的数学模拟。可以用微分方程等确定类数学模型进行模拟,也可以用随机过程等随机类数学模型进行模拟。例如用确定型模拟研究现有矿物相中元素的分布规律,以阐明岩浆过程和交代过程的特点。在随机过程模拟方面维斯杰列乌斯(A. Б. Вистелиус)作了大量工作,他建立了两种花岗岩数学模拟模型,一种为结晶花岗岩随机模型,另一种是经过交代改造的花岗岩随机模型,用这两种模型研究了加拿大渥太华附近的加齐诺岩体,苏联哈萨克斯坦的基捷尔塔斯岩体,西德的马利斯布尔克岩体。通过对几百个样品的实验研究表明,与高温热液矿床有成因联系的岩体和不成矿的岩体有某种随机特征的区别。当花岗岩本身就是分散元素矿石时,这种花岗岩的随机结构和不含矿岩体的结构截然不同。这一类研究可以导致建立全世界花岗闪长岩的概率模型。上述研究除了能解决地质找矿实际问题外,尚可解决一系列岩石学理论问题。

另外,通过数学模拟可以研究碎屑岩成因。碎屑岩的形成过程可以和一个数学过程相对应,即输入(相当于物质来源)→处理系统(相当于风化作用过程、侵蚀过程、搬运过程、堆积过程和成岩过程等地质作用过程)→输出(相当于最终产物:碎屑岩)。通过这种数学模拟对阐明碎屑岩成因取得了一定进展。

6. 在矿产地质工作中的应用

大致可分为以下四个方面。

(1) 在小比例尺区域矿产预测中的应用

应用数理统计方法可将一个研究区域分成有找矿价值区域和没有找矿价值区域,为布置找矿工作提供依据。美国哈里斯(Harris, 1966)把在一个地区内能找到矿床的概率和区域地质特征联系起来,建立了一种小比例尺区域矿产统计预测模型。他首先建立了矿产经济价值和地质变量之间的关系:

$$V = f_1(L, S, F, A)$$

其次建立了能够找到一个一定规模矿床的概率和地质变量以及矿产经济价值之间的关系:

$$P(v) = f_2(L, S, F, A, V)$$

上述二式中

V ——矿产的经济价值;

L ——岩性(包括岩石的时代和类型);

S ——构造形式;

F ——岩石裂隙特点;

A ——岩浆活动时代和接触关系;

$P(v)$ ——找到有一定经济价值矿床的概率。

为了把上述一般性的方程式化为可运算的形式,哈里斯在地质图上取了26个变量(如前寒武纪沉积岩所占百分数、侵入岩墙数目、不同类型构造数目等)。他将美国新墨西哥州和亚利桑那州的研究区域划分成243格,每格20平方英里,根据过去的开采记录确定了每一格中矿产的经济价值,然后用贝叶斯准则进行判别分析,确定出勘探的有远景地区和无远景地区。根据相同的思路,其他人提出了很多区域矿产统计预测模型,如主观概率模

型、统计决策模型，在实际应用中都取得了一定效果。

(2) 在矿产勘探阶段的应用

地质勘探工作（钻探、槽探、化探取样、物探测量）一般均按测线或网格布置。研究矿体形状和测线或网格密度之间的定量关系，可以为确定合理的测线或网格间距提供依据。

近年来，在矿产地质中开始应用数学探索理论（mathematical search theory）。例如，在铀矿勘探中，如果经过分析认为在勘探的第一阶段找到铀的可能性比第二阶段大三倍，则应用数学探索理论，可以提出在勘探的第一和第二阶段分配勘探手段的方案。如果对见矿概率估计正确，则上述分配方案比在勘探各阶段平均使用力量所取得的效果要好。

(3) 在矿产开发阶段的应用

近年来所采用的现代化开采技术，多以矿块为单元，进行大规模露天或地下开采。为此目的，要求对矿体有详尽的描述，要求建立综合考虑各种因素（矿体大小、形状、产地、储量、品位、矿石和废石比例）的全面数学模式。

为了计算不规则矿体的品位，需要把矿体分成许多立方体，对每个立方体进行品位计算，然后将所有立方体综合在一起，即得到整个矿体的品位数据。为上述计算过程编有专用的电子计算机程序，可以大大提高工作效率。

矿床品位和储量之间的定量关系也是一个重要的研究内容，一般认为两者之间有以下的指数关系

$$T = 10^{(10-G)}$$

式中， T 是吨数， G 是品位。如写成具有一般形式的公式则有

$$T = AB^{-G}$$

式中， B 为任意底数， A 是一个比例常数。

(4) 在矿产地质工作其他方面的应用

应用多元统计方法对矿床进行分类；区分具不同矿化类型的岩脉和不同矿物成分的岩脉。

在矿产地质工作中，获得的数据的数量通常是很大的，因此可以用电子计算机作地质数据的储存、索取和显示，建立专用的矿产地质数据文件或数据库。

7. 在石油地质中的应用

在初期阶段（五十年代）主要是将孔隙介质中液体流动理论用于多井的各向异性天然油贮中油、气、水流动的实际问题，计算出油田和天然气开采过程中油、气压力变化，并进一步评价油、气的可采程度。

六十年代以来，开始对地下油层数据作趋势面分析，并对地下油层作数学模拟。

数学方法和电子计算机在石油地质中另一重要应用是地质体的分类问题。如将含油的和不含油的地层加以区分，可以为石油勘探和开发提供有力依据。地质体分类方法主要为多元统计技术，如聚类分析、因子分析、判别分析等。

数学方法和电子计算机在石油地球物理工作中也得到了一定应用。如应用多元回归评价岩石参数；应用电子计算机自动和连续解释地球物理资料；应用多维相关研究地壳厚度，通过地质地球物理信息的统计分析来确定地壳结构；根据地质-地球物理资料研究局部构造；综合分析地质-地球物理指标并进一步预测局部构造含油性；根据地球物理资料

用多元回归方法, 或根据地表化探资料用因子分析方法发现含气层, 等等。

8. 在水文地质和工程地质中的应用

在地下水资源评价方面主要用最优化方法和数学模拟方法对地下水物理性质、化学和生物成分等资料进行系统分析, 可以精确地评价水量、水质, 为水量、水质控制和供水提供依据。

水文地球化学过程的数学模拟, 主要为根据一定规则建立地下水含水层中化学成分及其演化过程的数学表达式, 并进一步认识天然地下水中化学元素空间分布和迁移的规律。

在工程地质中主要为应用数学方法对工程地质信息进行加工, 建立综合的工程地质计算指标, 确定取样系统, 研究工程地质性质的空间变异等。

应用统计方法 (回归分析、相关分析), 可以建立一种指标和另外一些指标之间的相关关系, 用一些已知指标预测另一些指标, 例如有人对苏联列宁格勒附近的蓝色粘土建立了以下回归方程式

$$a_{2-6} = -0.0049 + 0.0068\epsilon + 0.038\epsilon^2$$

式中 a_{2-6} ——负载在 $2-6 \cdot 10^5$ 巴时的压实系数;

ϵ ——孔隙度系数。

另一回归方程为

$$E_0 = 7000 - 696.5W + 111.4\sigma_3 - 17.6W^2 - 5.5\sigma_3^2 + 58W\sigma_3$$

式中 E_0 ——形变模量;

W ——湿度;

σ_3 ——各个不同方向上的应力值。

根据以上方程可利用孔隙度系数、湿度、压力等指标来计算 E_0 , a_{2-6} 等指标。

应用回归分析和相关分析可以研究和预测工程地质过程。例如, 对于苏联伏尔加格勒水库岸边再造建立了回归方程式, 根据地质条件把水库库岸分成了三种类型: I 型、II 型、III 型, 并分别建立了回归方程式

$$\text{I 型} \quad Y = 1.17x_1 - 0.13x_2 - 0.056x_3 + 31.97x_4 - 5.02x_5 + 0.34x_6 + 54.24x_7 + 56.75$$

$$\text{II 型} \quad Y = 0.031x_1 + 0.033x_2 + 0.007x_3 + 10.79x_4 - 0.72x_5 + 0.075x_6 - 12.64x_7 + 8.85$$

$$\text{III 型} \quad Y = 0.59x_1 + 0.089x_2 - 0.038x_3 - 11.67x_4 - 0.595x_5 + 0.667x_6 + 7.72x_7 + 68.99$$

式中 Y ——1962—1971 年间岸边再造总量, 米;

x_1 ——1962—1971 年间总的波浪能, 吨·米;

x_2 ——1962—1971 年间高度为 0.5 米波浪的总的重复次数;

x_3 ——水位等于或高于正常设计水位的总天数 (1962—1971), 天;

x_4 ——均匀作用的波浪能和库岸夹角的多年平均值, 弧度;

x_5 ——浪侵阶地高度, 米;

x_6 ——浅滩坡度 (1962), 弧度;

x_7 ——其他变量。

应用回归分析方法研究滑坡速度和 12 种地质变量之间的关系, 建立了回归方程式, 可用于滑坡预测。

应用有限单元法、矢量分析法对岩体稳定性进行定量分析,可为边坡、隧道、矿坑的工程地质评价提供依据。

9. 在野外地质填图中的应用

电子计算机的自动储存和索取数据的性能,开始应用于野外地质填图工作。例如瑞典于1969—1970年间设计了一个野外数据的储存和索取系统(GEOMAP)。在野外填图工作中,GEOMAP系统采用标准野外数据卡片。卡片由几部分组成(如取样地点、参考依据、岩石特征、构造特征等),卡片上所填地质特征用数字或字母表示。地质工作者在野外工作中将观测到的地质现象按照要求以数字或字母的形式填写在卡片上,然后由一定的读取装置,将卡片上的记录读下来转换成磁带或穿孔卡片,直接送到电子计算机去处理。电子计算机配备有标准程序,可自动处理这些数据并绘出地质露头产状、岩石类型和构造特征的图件。野外数据卡片容纳不下的内容,可以记在“可处理记录”中,这样就使GEOMAP系统具有很大的灵活性。GEOMAP系统在连续几个野外工作季度中,经受了野外地质实际工作的检验,多数地质工作者认为这一系统的效果较好。

加拿大设计了几种野外地质数据系统,这些系统可以完成以下四方面的工作:

(1) 区域地质填图

在野外收集地质数据过程中,应用野外数据文件和野外数据卡片两种填写数据的格式。这些野外数据文件在每周末或每季度末校对一遍,并按穿孔的要求整理好,然后送往计算中心。野外数据以穿孔卡片或磁带形式输入计算机进行处理。野外文件由计算中心送回后,每200—250个文件装订成一个记录本,这些记录本可为地质填图通用并便于保存。野外地质数据经过初步校对、编辑和改正后,可以:(a)利用电子计算机标准程序算出统计分析结果;(b)利用XY绘图仪绘出任意比例尺的地质构造、岩性和其他有关图件。

(2) 地层记录和分析

设计了一套全部地层信息的标准记录系统(LEDA),可用于区域地质填图。这一系统的记录工作由7种卡片组成:卡片1——地点的详细描述,卡片2——地层单元描述,卡片3——岩性描述,等等。

(3) 地质钻孔的记录和分析

在加拿大目前有12个这样的系统,此外尚有专门用于石油钻孔数据的补充系统。这些系统可用于钻孔地质资料的岩相分析,也可利用钻孔地层数据进行快速地质编图。

(4) 矿床勘探

最近在加拿大设计出了一种RT系统,可用于处理各种地质数据,所采用的自由格式系统具有很大的灵活性。这种系统由文件、记录、数组、参数组成,每一个数据文件包括一系列记录,而记录包括一系列数组,数组又包括很多参数。这类数据系统成功地用于指导矿床勘探工作。

丹麦设计了一种称为SYSTEM-II的实验性数据处理系统。这一系统可用于处理野外数据和实验室数据。该系统首先用于格陵兰西部前寒武纪岩石地区的地质填图工作,以后又用于萤石的地球化学找矿,用于找镍矿,用于辉绿岩体的岩石学研究、花岗岩岩石物理性质研究等方面。

芬兰于1968年开始设计一种野外填图用的数据储存、索取和处理系统(GEOKU),并配备有自动查错程序。

四、数学地质工作中存在的主要问题

数学地质是一门新学科，二十年来得到了很大发展，成绩是主要的，但在前进中也有些缺点存在。

数学地质做为—门新学科，还有很多不成熟的地方，需要进一步摸索和探讨。

已有一些数学地质工作者，结合自己多年工作经验，综合苏联、欧美的数学地质文献，作了初步总结。这里仅就一些存在的主要问题谈一些初步看法。

(1) 有些数学地质研究课题是结合理论性的地质问题提出的，目前虽取得了一定进展，但尚没有造成地质学定量化研究方面的重大突破。

(2) 数学地质方法的研究最近不断取得新的进展，北美和西欧国家在这些研究中处于领先地位。但个别方法的突破对数学地质研究全局变化的影响不够明显。

(3) 在数学地质研究中地质工作者和数学工作者的结合尚不够。有些从事数学地质工作的地质工作者不善于提出重大的数学地质研究课题。数学工作者必须在数学领域进行创造性的研究，并深入地质实践，否则就不能在数学地质研究中发挥应有作用。

(4) 目前数学地质选题有自发性和盲目性的偏向，题目细小而分散，缺乏对长远性重大课题的规划和协调。

(5) 个别数学地质研究的地质基础工作不够坚实，因而影响了数学方法发挥更大作用。

(6) 有些数学地质工作的数学模型选择不合适，甚至有错用数学方法的个别事例。

(7) 个别数学地质研究课题中没有充分考虑数学方法所要求的前提和条件，因而所得结论不可靠。

(8) 对一些研究不充分的问题不宜匆忙作出结论，对作错误的结论要及时修改。如在数学地质的早期研究中曾得出以下结论：在矿物中稀散元素的富集分布规律具有对数正态性质，而在岩石中它的规律具有正态性质。这一结论是不对的，但没有及时修正，影响甚广，即是一例。

五、数学地质工作当前的主要任务及展望

对于数学地质工作当前的主要任务我们提出以下几点看法：

1. 进一步阐明数学地质的研究对象、研究内容和研究方法

任何—门成熟的学科都有自己明确的研究对象、研究内容和研究方法。由于数学地质是—门新学科，上述问题还不完全明确，需要通过研究来进一步阐明。

目前对上述问题已有一定研究。例如维斯捷列乌斯认为，地质学中的随机过程是数学地质的主要研究内容，概率论是最主要的数学工具。基于这种理解，他给数学地质下了定义：数学地质是关于地质过程的概念随机模型建立、检验和解释的科学。但上述看法有一定片面性。随机类数学虽然在地质学中占有重要地位，但在地质学中尚应用大量确定性数学方法以及其他数学方法。

加拿大数学地质工作者阿格特勃格 (F. A. Agterberg, 1974) 指出，数学地质是应用

数学方法研究地壳中问题的科学。这种提法太笼统。

在本文开头,作者提出了一种数学地质的定义,但这也仅仅是初步意见。

对上述问题的深入研究和进一步阐明有助于数学地质沿着健康的道路更快发展。

2. 进一步发展数学地质学科

在这方面有两大类任务。第一类任务是对发展比较成熟的数学地质分支,如地质学中的多元统计,应当进一步总结提高,深入探讨以下问题:地质多元统计在解决重大地质学理论问题和实践问题中的作用;多元统计方法在地质学中应用的前提、效果和应用条件;多元统计和地质学结合的一般规律等。第二类任务是促进数学地质中新分支(数学模拟、地质统计学、地质数据库)的进展。

3. 应用数学地质方法解决地质学中的重大理论问题

应用数学地质方法解决重大地质学理论问题是数学地质研究的一项重要任务,如构造地质问题、岩石圈构造动力学问题、岩石学问题等。

4. 应用数学地质方法解决重大地质实际问题

应用数学地质方法可能解决的地质实际问题很多,如应用数学方法预测矿产,通过物探数据处理说明地下地质结构和发现隐伏矿体,通过化探数据处理阐明背景值、异常值以及圈定勘探远景区等。

5. 探索地质学向定量化研究发展的科学的和有效的途径

地质学向定量化研究发展也就是地质学的数学化问题,无疑这是一个非常重要的问题。对这一问题发展的途径,存在着许多不同的观点,这是一项需要认真探索的带有根本性质的重大任务。

科学的数学化是现代科学发展的重要趋势,地质学也不例外。数学地质作为一门新学科的产生是地质学向数学化发展的必然产物。

由数学在地质学中应用萌芽的产生到数学地质学科的形成约有150年历史,其中最近20年是数学地质学科形成的关键时期,电子计算机在数学地质发展中起了重要作用。

数学地质作为一门新学科具有强大的生命力,今后数学地质将以更快速度和更大幅度向前发展,在解决地质学重大理论和实际问题方面将取得明显突破;地质工作者和数学工作者进一步的密切合作,将使数学地质发展到更高水平;数学和计算数学最新成果的不断引进,将使数学地质保持地质学中先进分支学科的特色。

我们相信,数学地质方法在广大地质人员中的普及,将推动地质科学向深度和广度更迅速发展,使找矿工作更富成效。

参 考 文 献

- [1] Agterberg, F. P., Geomathematics, Elsevier Scientific Publishing Company, New York, 1974.
- [2] Lantzy, R. J. et al., Catastrophe theory, Application to the Permian Mass Extinction, Geology, 12 (1977), 724—728.
- [3] Mc Cammon, R. B. (ed.), Concepts in Geostatistics, Springer-Verlag, New York, 1975.
- [4] Вистелиус, А. Б., Математическая геология—её основные направления и задачи, Советская Геология, 1(1977), 11—34.
- [5] Воронин, Ю. А., Методические вопросы применения математических методов в геологии, Новосибирск, 1974.

聚类分析的方法及其应用

方 开 泰

聚类分析 (cluster analysis) 是研究物以类聚的一种方法, 在地质界有人称它为群分析、点群分析、簇群分析等。在 1975 年概率统计会议上 (苏州) 已商定采用聚类分析作为它的名称。

人类认识世界往往首先将被认识的对象进行分类, 因此分类学便成了人类认识世界的基础科学。在古老的分类学中, 人们主要靠经验和专业知识实现分类。随着生产技术和科学的发展, 分类越来越细, 要求越来越高, 有时只凭经验和专业知识不能进行确切分类, 于是数学这个有用的工具就逐渐被引进到分类学中, 形成了数值分类学, [3][4][5] 是这方面的代表作。近十几年来, 数理统计的多元分析方法发展十分迅速, 多元分析的技术自然地引进了分类学, 于是从数值分类学中逐渐地分离出了聚类分析这个新的分枝, [1][2][6][7][8] 对聚类分析作了系统介绍。

聚类分析已广泛用于地质、气象、水文、考古、医学、心理学等方面。大量事实证明, 聚类分析是一种行之有效的办法。近年来, 在我国地质界应用聚类分析方法者日益增多, 为了使地质工作者对聚类分析有一个比较全面的了解, 本文将主要的方面作一梗概介绍。

§ 1、变量和标本

在地质工作中有时需要将标本进行分类, 这种分类称为 Q 型的, 有时要将变量分类, 称为 R 型的。为了进行分类, 首先就要研究变量和标本之间的关系。

设有 m 个变量, $X_1, \dots, X_m$; n 个标本, $x_1, \dots, x_n$ 。每个标本都是一个 m 维向量。第 i 个标本的第 j 个变量以 x_{ij} 表示。描述变量和标本之间的关系常用如下的方法:

1. 距离

将标本看成 m 维空间的一个点, 可定义标本之间的距离, 距离相近的标本归成一类, 距离远的标本应属不同的类。标本 x_i 和标本 x_j 间的距离最常用的是绝对值距离 $d_{ij}(1)$ 、欧氏距离 $d_{ij}(2)$ 。切比雪夫距离 $d_{ij}(\infty)$

$$d_{ij}(1) = \sum_{k=1}^m |x_{ik} - x_{jk}|, \quad d_{ij}(2) = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2} \quad (1.1)$$

$$d_{ij}(\infty) = \max_{1 \leq k \leq m} |x_{ik} - x_{jk}| \quad (1.2)$$

它们分别是明考斯基距离 $d_{ij}(q)$ 当 $q=1, 2, \infty$ 时的特例,