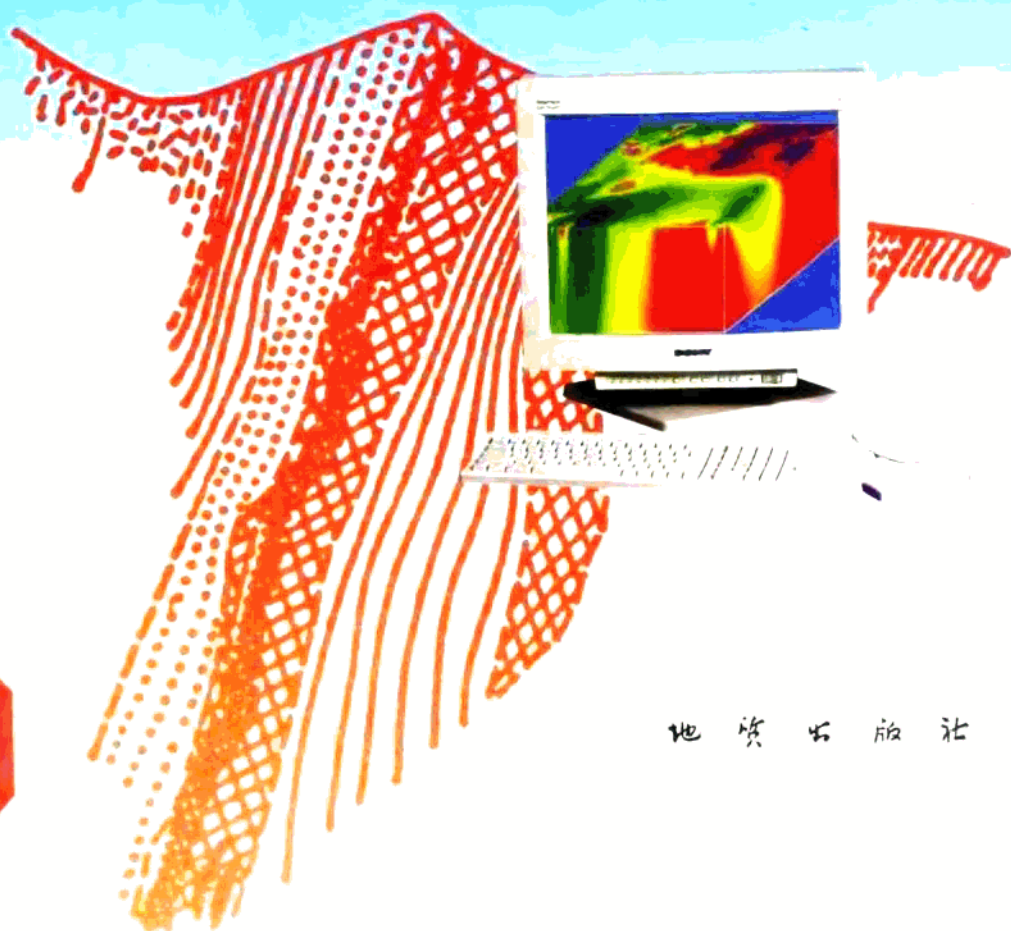


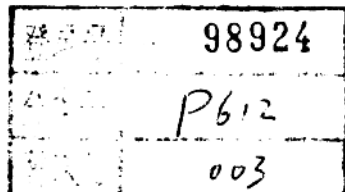
成矿预测方法通则之二

成矿预测方法

朱裕生 肖克炎 等 编著



地质出版社



200845684

成矿预测方法通则之二



成 矿 预 测 方 法

540627

朱裕生	肖克炎	丁鹏飞	金丕兴	李纯杰
梅燕雄	曾昭健	黎世美	瞿伦全	陈圆明
	宋玉玖	江西根	毕惠亭	编著



00927215



地 质 出 版 社

· 北 京 ·

内 容 提 要

作者通过十多年的实践总结了一套成矿预测的有效方法,是目前我国较系统的预测方法专著之一。书中重点介绍了成矿预测的方法体系、成矿预测系列图件的编制、成矿预测工作的实施、成矿预测的技术要求和成矿预测的系统实例以及找矿靶区的优选方法。提出成矿预测的内涵是将地质工作获取的多元地质数据转化为矿产资源信息,用以指导矿产勘查。本书详细地展示了经矿产勘查获得实效的范例,以促进成矿预测工作的发展。

本书可供矿产勘查人员阅读,也可供科研和教学人员参考使用。

图书在版编目(CIP)数据

成矿预测方法/朱裕生等编著. -北京:地质出版社,1997.4(成矿预测方法通则;2)

ISBN 7-116-02298-8

I. 成… II. 朱… III. 成矿预测-方法 N. P612

中国版本图书馆CIP数据核字(96)第23123号

地质出版社出版发行

(100083 北京海淀区学院路29号)

责任编辑:杨友爱

责任校对:梁毅

*

北京地质印刷厂印刷 新华书店总店科技发行所经销

开本:787×1092 1/16 印张:21 字数:500000

1997年4月北京第一版·1997年4月北京第一次印刷

印数:1—500册 定价:37.80元

ISBN 7-116-02298-8

P·1725

前 言

以矿床和矿体为主要对象的成矿预测，是一项贯穿勘查全过程的工作，即从基础地质工作开始（甚至包括与矿产地质研究有关的科研工作）直到矿床勘探和矿产开采，都在进行相应阶段不同要求和不同比例尺的成矿预测工作^①。当前，地质找矿已从地表露头的寻找转入隐、盲矿床的预测、普查和评价；地质工作转入普查的途径发生了根本的变化；在地质找矿的投入（资金和工作量）逐年减少的情况下，为争取地质找矿的重大突破、深化新一轮固体矿产普查，在成矿远景区内筛选“重中之重”的普查效益倍增投入区，达到寻找和评价大型及超大型矿床的目的，获取地质找矿的最大效益，提出加强成矿预测工作。

成矿预测是应用地质理论和科学的方法综合地质、地球物理、地球化学和遥感地质等基础地质工作所获得的地质-找矿信息，总结成矿地质条件和矿床赋存的规律，建立矿床模式，圈定不同级别的成矿区带、矿田、矿区及矿床范围内不同类别的预测区或三维空间的找矿靶区，提出开展矿产勘查的重点区和勘查工程施工的普查“点”，指导不同层次的地质找矿工作的实施，提高地质找矿的科学性、有效性，实现重大突破的目的。

成矿预测的基本任务是将地质工作获取的数据转化为矿产资源的信息（又称提取成矿信息），缩小找矿地工作区、在优化基础上，提出最佳普查区。这项任务的实现，除了应用已有的基础地质和成矿学的理论外，关键是选择合理的方法。将地质数据转化为矿产资源信息的整个过程是成矿预测方法使用的过程（包括预测人员的思维方法和经验）。

成矿预测方法是一个体系，用单一的方法是不能实现预测的，也达不到目的。从成矿预测方法的体系出发，才能推断将来可能或应当发现的潜在矿床，获得较可靠的预测成果。从这一全过程出发，成矿预测是成矿信息提取、信息综合、成矿理论应用、找矿靶区圈定和优化、工程验证的总体结果。成矿预测方法贯穿在预测过程的每个阶段和步骤实施的具体工作中，称之为“预测方法体系”。

在成矿预测工作中，成矿预测人员要重视方法的选择、应用和效果的验证。本书阐述成矿预测方法体系，力求“体系”和内容相结合，供广大成矿预测人员参考应用，并按自己预测工作的实施，辟择新的途径，充实和完善这一“体系”。

本书的第一、五章由朱裕生编写；第二章由朱裕生、丁鹏飞、陈圆明、宋玉玖编写；第三章由朱裕生、肖克炎、金丕兴、李纯杰编写；第四章第一、二节由毕惠亭、江西根等提供资料和报告，由丁鹏飞、陈圆明、宋玉玖、梅燕雄编写，第三节由曾昭健编写。

^① 见中、大比例尺成矿预测文献汇编中的“成矿预测的地质基础和方法”，朱裕生编著，1991，P.43—158。

目 录

前 言

第一章 概述	(1)
第一节 成矿预测方法体系略述.....	(1)
第二节 数据.....	(4)
第三节 单元	(12)
第四节 控制区 (已知区)	(13)
第二章 成矿预测方法	(15)
第一节 经验类比法	(16)
第二节 地质-地球物理法	(19)
第三节 地质-地球化学法	(50)
第四节 综合技术方法	(65)
第五节 找矿模型	(69)
第六节 成矿预测系列图件编制方法	(94)
第三章 找矿靶区优选方法	(106)
第一节 概述.....	(106)
第二节 找矿靶区优化的基本问题.....	(107)
第三节 找矿靶区优选的实施.....	(113)
第四节 找矿靶区优选的方法.....	(117)
第四章 成矿预测方法实例	(160)
第一节 山东变质岩区 (水道—冯家地区) 的成矿预测.....	(160)
第二节 香花岭地区的综合方法成矿预测.....	(175)
第三节 湖南郴桂地区中大比例尺成矿预测方法及成效.....	(221)
第五章 成矿预测工作的实施	(259)
第一节 基本情况.....	(259)
第二节 实施中的几个关键问题.....	(263)
附件:	
附件一 成矿远景区划基本要求 (试行), 1980.2	(275)
附件二 矿产资源总量预测试行基本要求, 1983.8	(279)
附件三 固体矿产成矿预测基本要求, 1990.7	(291)
附件四 固体矿产第二轮成矿远景区划技术要求 (试行), 1992.11	(304)
附件五 固体矿产第二轮成矿远景区划成果汇总要求, 1993.8	(320)

第一章 概 述

第一节 成矿预测方法体系略述

成矿预测方法体系的内容包括以下几方面：

1. 对原始数据进行收集、汇总、编辑和制卡；
2. 对数据进行检验、选择、转换、补齐，对有些描述性的数据必须重新取值；
3. 按预测的目的、任务划分已知区或确定重点研究的代表性矿床，确定已知区；
4. 根据成矿地质理论和研究方法，分析成矿地质背景，研究成矿规律，总结控矿因素，归纳找矿标志；
5. 选用当前有效的建模理论和方法建立矿床模式和找矿模型；
6. 按成矿预测的基本理论、原则和有关规定划分成矿区带，圈定预测区；
7. 选用适用的成矿预测方法作定性和定量预测，优选和确定预测区的级别；
8. 编制成矿系列图件。

成矿预测方法体系充分体现了由地质数据转换成潜在矿床过程中所应用的各种方法的全过程。在当前开展的成矿预测工作中，需要对其中的重点问题作出论述。

一、定性和定量预测问题

在以中、大比例尺为核心的成矿预测工作中，定性预测是主要的，它包括研究成矿规律、建立模式（矿床模式和找矿模型）、划分成矿区带和成矿远景区（或找矿靶区）、靶区优选和标定级别。定量预测仅对潜在资源量作出估算，定量预测根据相似类比理论阐明潜在矿床的几何特征是矿体的规模、产状、形态、空间位置、埋深、边界等矿床特征的描述问题，它是预测工作中的主体，它仅对上述描述内容作出具体阐述；定性预测的结果指导验证工作的布置，而定量预测目前使用的种种方法仅仅是量的概念，它若是在定性预测的基础上作出定量估算则是具有实际意义的。

定量预测与成矿预测工作中使用有效的数据处理方法是有原则区别的。采用有效的方法对地质数据进行处理，以揭示地质体的各种属性的统计规律，如各种特征值、平均值、标准差、方差、协方差、相关系数、变化序列、偏倚系数、峰凸系数、复相关系数、混合分布类型、总体分解后各总体的比例等都具有实用价值，从中总结地质规律，是成矿预测中必要的工作内容，但它不属于定量预测工作范畴，最多仅属于定量预测的基础工作。有些数据处理是不同地质学分支的专业基础工作，例如化探数据的处理、重磁数据的处理、航卫资料数据的处理等，它们是为了解决圈定异常、消除干扰等实际问题所作的必要工作，同样不属于成矿预测的方法；有些数据处理，像矿与非矿信息的区分、成矿信息的浓缩、信息综合等仅属于成矿预测方法的外延部分，同样不能算是定量预测的内容。严格地讲，定量预测是应用经验的和数学的（以统计学为主）的方法对潜在资源量作估算的全部内容。据

此可知，定量预测的内容仅在成矿预测方法体系中占较小的份额。加之，当前地质理论尚处于经验综合和感性认识阶段，从总体上看仍属于经验性的工作，有些描述内容仍处于“模糊”阶段，而定量预测方法是建立在数学模型基础上的，它具有严格的逻辑关系，应用数学模型估算的矿产资源量带有较大的风险。基于上述种种原因，成矿预测必需以定性为主，定量预测成果只能在圈定和优化后的找矿靶区范围内进行，才具有一定的实用价值。

二、中大比例尺成矿预测的基本特点

如前所述，成矿预测分小、中、大三个层次，每个层次的目的、任务是不同的，使用的成矿预测方法也各异。这种差异反映到成矿预测的各个工作步骤中，现用表 1—1 来说明。

表 1—1 不同层次成矿预测之间的差异

顺序	类 目	小比例尺	中比例尺	大比例尺
1	圈定的成矿区带级别	I—II (特殊情况下 N 级)	IV—V	找矿靶区
2	目的	圈定成矿远景区	矿田预测区	矿床、矿体
3	使用的资料	已有的地质勘查资料	有关的 1:20 万地、物、化、遥感资料	与预测比例尺配套的地、物、化、航片资料
4	预测方法	地质学和数学地质方法	经验的、地质学的和数学地质的方法	模式类比、理论预测和综合技术方法
5	对预测成果要求	定性加定量 (G 级)	定性为主，定量达 F 或 G 级	E+D 级及四定 (定规模、定位、定量、定类型)
6	成果作用	部、局作中、长期规划	局、队部署矿产勘查	队 (或局) 普查选区依据
7	工作承担单位	生产或科研单位	以生产单位为主	基层地质队、科研配合
8	验证要求	间接验证	间接和直接验证	提出山地工程直接验证的位置，为普查设计提供依据

小比例尺成矿预测大体上与 1979 年开展的第一轮区划和跨省区划相当；中比例尺成矿预测相当于当前正在开展的第二轮成矿远景区划；大比例尺成矿预测是当前重点找矿的一项工作。

从表 1—1 可知，三者的差异是很明显的，特别是预测的目的、预测要求和使用的资料三方面的差异较大，因而使用的成矿预测方法也各不相同。在中、小比例尺的成矿预测工作中，对使用的预测方法要求不是很严格，而大比例尺成矿预测方法其要求较为严格，对一些可靠性较差的预测方法（如丰度法）很少使用，而侧重应用地质人员在矿产勘查中积累的经验，常常采用经验类比法。对隐、盲矿床的预测来说，常常使用综合技术方法进行推断解释。

小比例尺成矿预测侧重于成矿区带的划分和成矿远景区的圈定；中比例尺成矿预测侧重于矿田预测区内有无潜在矿床的推断解释；大比例尺成矿预测侧重于矿床三度空间的变

化和几何形态的研究和推断解释,不断向立体空间的判断延伸。总之,大比例尺成矿预测的这些基本特点决定了成矿预测方法的选择和要求应有以下的特殊性:

1. 定性预测是中大比例尺成矿预测的主体,目前的定量预测使用的数学方法(统计学)在中、大比例尺,特别在大比例尺预测研究区范围内,限于已知矿床过少而无法建立预测模型;

2. 大比例尺成矿预测对矿体(床)特征和形态的研究探索工作是预测的核心,不断向三度空间延伸,它常常用建立矿床模式和找矿模型的途径来实施,模式类比的预测方法是比例尺预测的主要方法;

3. 圈定的找矿靶区要具有过硬的资料依据,预测效果将直接受普查工程验证的检验,所以它必须提交可靠性较高的预测成果。

综上所述,大比例尺成矿预测要求高、难度大,是与普查找矿相衔接的工作。

三、大比例尺成矿预测的选区和必备条件

大比例尺成矿预测能否取得成效,取决于选区是否正确。在通常情况下,大比例尺成矿预测区的选择,要通过专门的系统工作步骤的探索和论证。按成矿预测的程序,大比例尺成矿预测选区应在V级成矿区或中比例尺成矿预测所圈定的A类预测区内进行,也可在已知矿床的深部或其外围进行。选择大比例尺成矿预测工作区的必备条件有以下几方面:

1. 工作区内作过1:200000中比例尺成矿预测,在已圈定的A类或在1:50000区调的工作区范围内作为选区对象。

2. 必备的基本资料条件:

- ①与预测比例尺相应的地质矿产图(矿田成矿区或矿床);
- ②与预测比例尺相应的地磁资料;
- ③与预测比例尺相应的重力资料;
- ④与预测比例尺相应的电法资料(激电、自电或其它电法);
- ⑤与预测比例尺相应的化探资料(次生晕或原生晕);

3. 具有可供类比的已知矿床的综合研究成果:

- ①成矿模式;
- ②找矿模型;
- ③已知典型矿床的地、物、化综合剖面。

它是相似类比的标准。

4. 预测的目标任务要明确。

大比例尺成矿预测具有多学科、多结构、多层次的特点。多学科系指地质、地球物理、地球化学、重砂、遥感、数学地质方法和计算机技术等。多结构指预测方法,它需野外调查和室内研究相结合,直接观察和间接观察相结合,理论和实际相结合,点和面相结合,多结构的工作方法贯穿在每个工作步骤中。多层次系指普查勘探阶段不同、预测的目的任务也不尽相同,大比例尺又分矿田和矿床预测两个层次。矿田预测系指1:50000比例尺的预测;矿床预测系指1:10000或更大比例尺的预测,其中的1:5000和1:2000比例尺的预测属勘探阶段预测矿体的任务。对不同的目的任务,预测方法不能相互混淆;否则,造成目标任务不明确,致使预测过程中造成混乱,达不到预测的目的。

5. 要有完成预测任务的必要人材。

从多学科出发,它要具有掌握地、物、化、航卫知识的综合人材和各技术学科的专门人材。在工作的实施过程中实行以下三种结合:

- ①地、物、化、遥感专业的结合;
- ②以地质队为主,实行地质、科研和教育的三结合;
- ③地、物、化专业人员在工作中的结合。

具备上述各项基本条件后,在预测过程中,以地质成矿理论为基础,实现以地质找矿的突破为目的,以各种先进的数据处理方法和成矿预测方法为手段,逐步提高找矿靶区的类别,不断缩小找矿区的面积,提出验证工程的正确位置,达到发现矿床的目的。

四、成矿预测的重点目标——大型、特大型矿床的预测^[3]

大型、特大型矿床的预测和发现不属于某种预测比例尺的任务,各种比例尺的成矿预测都可能预测和发现大型、特大型矿床。但在成矿预测的整个过程中,要始终抓住大型、特大型矿床的这个大目标进行预测评价。根据当前积累的有限经验,有以下的途径进行预测评价:

1. 预测工作中,从大型、特大型矿床的角度出发进行评价,如玉龙斑岩型铜铅矿床。
2. 在预测评价时,由于已有的地、物、化、遥感资料尚未显示大型、特大型矿床的信息,所以在预测普查过程中应逐步充实这类信息,使得它从中、小型矿床勘查开始逐步增大,最后达到大型或特大型矿床,如安徽铜陵的狮子山矿田中的冬瓜山矿床,是在勘查中、小矿床过程中发现和扩大到大型矿床的规模。
3. 由于工业上工艺流程的改变,新技术的应用,过去认为是表外矿,不属于矿床的那一部分地质体而成为新的矿物原料。如有些高岭土矿床,随着加工技术的改进,原来不认为属涂布、括刀级高岭土而现在上升到此类级别,且具相应规模。这样的实际例子我国已有几处。
4. 新类型矿床的发现有可能属大型、特大型矿床者。

这些途径,在成矿预测中要时时把握,争取预测和发现大型、特大型矿床。预测和寻找大型、特大型矿床的宗旨应贯穿在成矿预测的整个过程中。

第二节 数 据

数据是成矿预测的基础,没有数据,也就没有成矿预测,无论是传统预测也好,近代的定量预测也好,都是这样。在地质学中,特别是在成矿预测工作中,数据包括地质、地球物理、地球化学和航卫四方面的内容。可以说,在成矿预测工作中,数据的内涵是广泛的。总的来说,表示某一地质现象所取的不同数值的量(定性和定量的)都称地质数据。

当前,在成矿预测工作中使用的数据,其来源是多种多样的,有的是用先进的手段和方法测试的,如航磁、化探扫面数据等;有些是地质人员的描述数据,如野外填图时对岩石的描述及岩矿定名等;有些数据是在室内综合研究后根据地质理论按预测方法要求赋与的定性数据。从成矿预测的需要和应用出发,对数据作如下几方面的叙述。

一、数据分类

对于地质数据,从不同的角度出发,可作不同的分类。习惯上一般易将地质资料和地质数据混淆起来。我们所说的资料是指数值的、文字的和图件的;地质数据是指地质生产

和研究过程中所需要的那部分资料,又常常分为成矿预测数据、同位素数据、岩石化学数据、古生物数据等;地质变量是指参与某种需要而建立数学模型的成分和参数。地质变量来源于地质数据,地质数据又源于地质资料,据此对地质数据可作如下分类(图1—1):

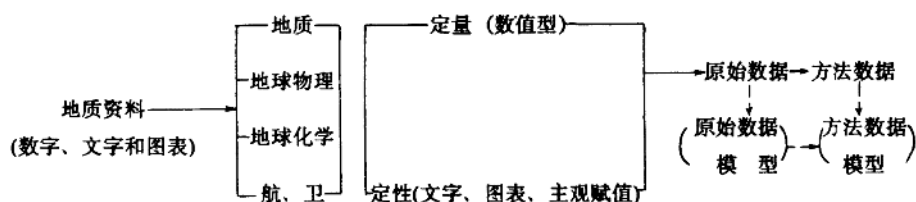


图1—1 地质数据的分类图

图1—1舍去了对数学地质的要求和数据统计特征(如连续型、离散型、定和型、名义型、有序型等)的要求,是以成矿预测的需要作的原则分类。

成矿预测使用的数据水平是有差异的,有些是精度达 10^{-9} 级的量值,如金元素;有的是达 10^{-6} 级的值,如大多数铜、铅、锌等有色金属等;有些是经验数据;有些是按地质理论赋予的定性数据(如有矿为1,无矿为0等)。而用这些数据进行处理,目前总的来说,还未达到系统化程度,可以说相当落后,尚未把其中的成矿信息提取出来。现有地质数据在内容上和空间上都是随意的,在大多数情况下,成矿预测人员都在被动地迁就当前地质工作的现状,将原来随意的数据勉强组成方法数据模型。这种随意的原始数据模型将影响成矿预测的结果,甚至造成预测的失败。因此,在成矿预测工作中要解决使用什么数据,怎样使用数据的现实问题。使用数据的总目标是设法构造一个与地质矿产的实际吻合程度较高的数据模型。

原始数据模型是使用地质、物探、化探和航卫数据获得的,但这些数据是按各地球科学分支学科的需要和要求获取的。它们对各自的专业来说是合理的,但对成矿预测工作来说,有些是合理的,有些不尽合理,有些无法使用。选用地、物、化、航卫资料,并非一切资料都能选取而组成原始数据模型。它要按预测的目的、任务和选用的方法要求,使用合格的数据组成原始数据而构置成原始数据模型。

从严格意义上讲,按照成矿预测数据处理要求和提取成矿信息(在本书第二章中详述)的需要,编制一个综合各类资源的原始数据格式,它基本上要囊括描述地质学和技术方法学科的全部内容。建立这样的数据基础后,再按数据处理的需要产生其它的数据模型,这就是原始数据模型。

组成原始数据模型中的各类数据都受到当时条件的限制,它与成矿预测的要求(系统、全面、可靠)相差较大,通常是零星的、片断的,由不同单位和不同人员提供的,数据与数据之间的可靠性相差甚远。在这种情况下,对原始数据进行检验、取舍是必要的。此项工作的完成与否,决定了原始数据模型的优劣,它需要由丰富经验的地质人员和成矿预测人员共同努力完成。

方法数据模型是在原始数据模型基础上生成的。众所周知,每种成矿预测方法,对数据模型往往都有特殊的要求,若不满足这种要求,则由方法本身产生的不确定性(例如有

些统计学方法存在有偏估计)会给预测结果带来干扰。例如回归分析的原始数据都要通过正态检验,若不服从正态分布,则由计算带来的误差干扰无法评估,其结果的可信度中包含了这种干扰而降低预报的可靠性。由原始数据模型转化为方法数据模型,通过数据转换,对原始数据模型中的某项或某几项数据进行变换,它可以是元素比值,也可以是相对熵、取对数、相关系数、因子得分等构置的新变量。由这些新变量(符合方法处理要求的)构成的数据模型称为方法数据模型。

在当前使用计算机的情况下,构置方法数据模型就容易得多。在明确了使用的成矿预测方法后,按方法要求,在计算机上比较容易地构置成方法数据模型。

在成矿预测工作中,对地质资料作上述的处理和确定各个具体步骤的要求是较为重要的,否则将会导致成矿预测工作的失败。

二、数据补齐

在原始数据模型中,有时出现由于原始资料的不齐而缺失某几个数据,尤其是某些与成矿关系较为密切的数据项中缺失几个。在这种情况下,需要找到一个合理的方法,将缺失的数据补上,以弥补它的空缺。补缺数据必需在不增加或不丢失成矿信息的前提下进行。目前常用的有:①插值法;②平均值替代;③统计法,用平均值与方差的相互关系作出取值的判断;④主观法,地质人员按成矿规律作出补缺数据的取值;⑤趋势外推法;⑥滑动窗法。

对某些空间数据来说,由于各区段的地质工作程度存在差异,存在数据过密及过稀的问题,也会对预测结果造成影响。通常的做法是在数据过密的区段合理地抛弃过密的样品,在过稀的区段需要使用“补救”的办法使样品数量达到基本合理,构置一个合格的方法数据模型。目前,在使用化探扫面资料时,若个别样品缺失,则使用插补或外推法补齐;在航磁中,过密地段采用抽线的方法处理,而过稀的地段,则采用先圈定等值线,再取数,或直接在等值线图上取值的方法。但是目前对数据“补齐”方法研究甚少,对此尚要进一步探索研究。

三、数据转换

由原始数据模型转换为方法数据模型时,数据转换是一项不可缺少的工作内容。

现有的地质数据的类型不同,量纲不一,数据之间的关系各异(如线性、非线性等)。故以原始数据模型出现的数据集在大多数情况下无法用现有的数据处理方法和预测方法的程序进行处理,因为它不满足数学模型的要求。在这种情况下,成矿预测人员从理论上要补齐“数据”,而不是“变量”,程序处理的要求是地质变量。此时,必须对原始数据模型的逐个变量进行研究,采用数据转换的方法予以完成。数据转换的根本目的是构置方法数据模型,但在每一过程中都有明确而具体目的:

1. 使地质数据尽可能呈正态分布;
2. 统一地质数据的量纲;
3. 使地质数据的非线性相关关系变换为近于线性相关;
4. 用一组新的少的而又独立的地质数据代替原来一组有不同相关关系的数据。

从地质数据变换过程中不损失信息、不增加干扰的原则出发,其数学逻辑关系是:地质数据的变换值是观察值的一个函数,它定义了一个新的观察值,这一点可以用下列的函数表达式表示^[4]:

$$U = f(X) \quad (1-1)$$

式中: X 为原始观察值; f 为变换的函数形式; U 为新观察值。

不论地质数据变换的方法多么不同和多么复杂,它都可以用公式 1-1 的形式来表示。变换后所获得的新数据值,可以是一个简单的变量,也可以是一个向量,甚至是一个矩阵。通过变换,我们期望获得的新变量能达到上述的一个或几个目的,特别是使变换后的变量符合评价方法对方法数据模型的要求。由于目前已有各种方法可供评价人员选择,这里把主要方法加以扼要叙述。

1. 网格化变换。在评价时,对有些原始数据需要作网格化变换,公式 (1-1) 便改变成下列形式:

$$U'_i = \frac{\sum_{i=1}^n X_i}{n} \quad (1-2)$$

式中: U'_i 为网格比值, $i=1, 2, \dots, n$; X_i 为单元内(网格内)某地质变量的 i 个观察值; n 为单元内的观察次数。

网格化变换是矿产资源评价中对地质数据最经常的和最普通的一种变换方法。在矿产资源方法学中往往把网格与单元统一起来。

2. 标准化变换

为了使原始地质观察数据的量纲一致,需要对原始数据进行标准化。经标准化变换后获得的新变量 U'_{ij} 的均值为 0, 方差为 1。各变量的每个观察值处于同一量纲,而且变换前后和其它变量之间的相关程度不变。标准化变换时,公式 (1-1) 便变成下列形式:

$$U'_{ij} = \frac{X_{ij} - X_i}{s_i} \quad (1-3)$$

式中: U'_{ij} 为变换后获得的第 j 个变量的第 i 次的新观察值, $i=1, 2, \dots, n$, $j=1, 2, 3, \dots, p$; X_{ij} 为第 j 个变量在第 i 次的原始观察值; X_i 为 i 个数据的算术平均值; s_i 为 i 个数据的标准差。

经标准化后获得的新观察值 U'_{ij} 之间及任一 U'_{ij} 的值与 U_j 之间的相关程度不变。例如,二个数据的原始观察值如表 1-2 所示,标准化后的新观察值列于表 1-3。

表 1-2

X_{ij}	0.5	1	1.5	2.15	2.75	3.25	3.75	4	4.5	5	...
U'_{ij}	0.3	0.65	0.8	1.05	1.3	1.6	1.85	2.07	2.11	2.4	...

表 1-3

X'_{ij}	-1.01	-0.79	-0.58	-0.30	-0.04	0.18	0.39	0.50	0.72	0.93	...
U''_{ij}	-1.84	-1.26	-1.01	-0.60	-0.19	0.31	0.73	1.09	1.14	1.64	...

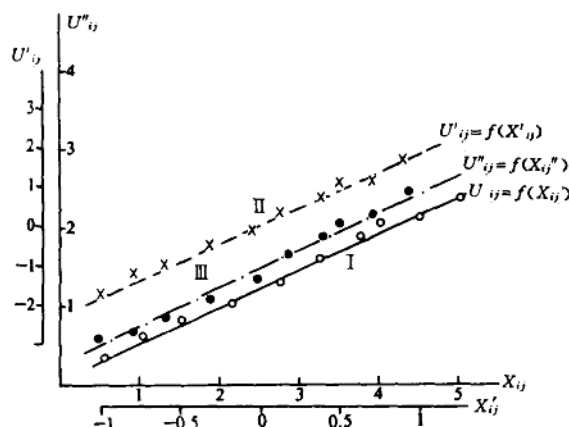


图 1-2 数据标准化前后的相关关系图示

根据表 1-2 和 1-3 作其各自的点聚图 (图 1-2)。从图 1-2 中明显看出, 直线 I 表示原始观察值之间呈线性相关, 将相应数据标准化后 (直线 II), 它们仍然呈线性相关。若将 X_{ij} 标准化, U_{ij} 仍用原始观察数据时, 则曲线 III 仍呈线性相关。这一点, 应用概率论的基本原理也可以从数学上获得证明 (证明从略)。数据转换后其相关程度不变的原理, 是在成矿预测工作中对原始数据实行转换的理论依据, 否则将会影响到方法数据模型的合理性。

3. 极差变换

在数据处理中, 当观察值的量纲和数值差异太大时需要将原始数据进行极差变换 (亦称对数化或正规化)。用极差变换后获得的新数据在 0 与 1 之间, 而变换前后和其它数据的相关性不变。极差变换的方法是在每个变量中挑选最大值和最小值, 这时的公式 (1-1) 就成为下列形式:

$$U'_{ij} = \frac{X_{ij} - X_{j\min}}{X_{j\max} - X_{j\min}} \quad (1-4)$$

式中: U'_{ij} 为变换后获得的第 j 个变量的第 i 次的新观察值, $i=1, 2, \dots, n, j=1, 2, \dots, p$; X_{ij} 为第 i 个观察点上第 j 个变量的原始观察值; $X_{j\min}$ 为第 j 个变量数据的最小值; $X_{j\max}$ 为第 j 个变量数据的最大值。

4. 对数变换

当数据呈偏态分布时 (如地球化学测量的各元素的浓度值), 需要应用其它的方法变换 (对数变换、平方根变换、反余弦变换、反正弦变换等), 变换后的新观察值将接近于正态分布。

当 n 个数据中, 其低值多, 高值少, 则这 n 个数值为正偏分布, 且正偏的偏度大, 可采用对数变换, 这时公式 1-1 成为下列形式:

$$U'_{ij} = \ln(X_{ij}) \quad (1-5)$$

其中 U'_{ij} 和 X_{ij} 的意义同公式 (1-4)。

这样获得的新数据值的分布将得到改善，甚至接近正态分布。

其它像反正弦、反余弦、平方根变换的原理与对数变换的原理类似，都属于偏态变为正态的变换方法。具体选择哪种方法，视原始数据的统计特征而定，属负偏者用反正弦变换，属正偏度大者选用对数变换，偏度中等选用平方根变换，属弱正偏用反余弦变换。这些都是原则上的叙述，在实际工作中，对原始数据要进行统计描述。

5. 线性变换

线性变换的方法多种多样，一般是选择某种图像的方法去拟合原始数据在平面上的分布趋势，然后用该图像的方程进行适当的变换。本文选用简单的常见的几个函数图像作示范，供建立方法数据模型的人员参考。

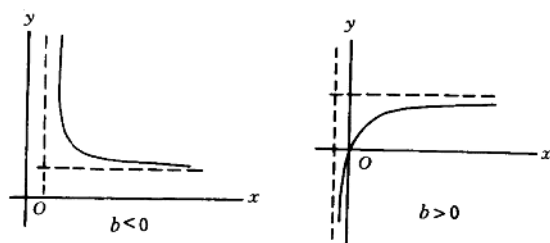


图 1—3 双曲线函数图像

$$(1) \text{ 双曲线 } \frac{1}{y} = a + \frac{b}{x} \quad (\text{图像见图 1—3}) \quad (1-6)$$

$$\text{变换时, 令 } y' = \frac{1}{y}, \quad x' = \frac{1}{x},$$

$$\text{则有 } y' = a + bx' \quad (1-7)$$

公式 (1—7) 是公式 (1—6) 的线性变换形式，即由双曲线方程 (见图 1—3) 变换成回归方程。

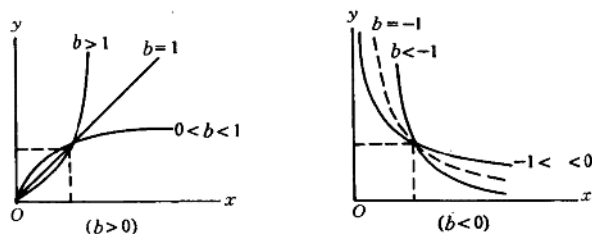


图 1—4 幂函数图像

$$(2) \text{ 幂函数 } y = dx^b \quad (\text{图像见图 1—4})$$

变换时两边取常用对数得：

$$\lg y = \lg d + b \lg x$$

$$\text{令: } y' = \lg y, x' = \lg x, a = \lg d$$

$$\text{则 } y' = a + bx' \quad (1-8)$$

致使公式 $y = dx^b$ 变换成线性回归形式。

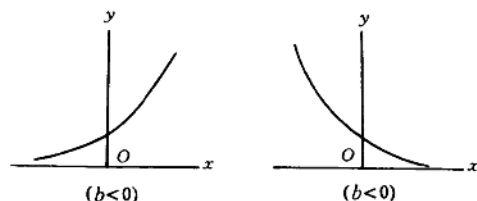


图 1-5 指数函数的图像

(3) 指数函数

$$y = de^{bx} \quad (\text{图像见图 1-5}) \quad (1-9)$$

对它的变换两边只要取自然对数，得：

$$\ln y = \ln d + bx$$

$$\text{令 } y' = \ln y, a = \ln d,$$

$$\text{则 } y' = a + bx \quad (1-10)$$

公式 (1-9) 变换成公式 (1-10) 的线性回归方程。

(4) 对数函数

$$y = a + b \lg x \quad (\text{图像见图 1-6}) \quad (1-11)$$

变换时令 $x' = \lg x$

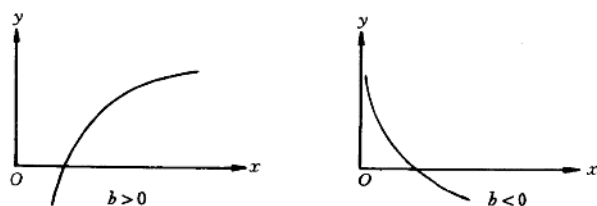


图 1-6 对数函数的图像

$$\text{则有 } y = a + bx' \quad (1-12)$$

$$(5) y = de^{\frac{b}{x}} \quad (\text{图像见图 1-7}) \quad (1-13)$$

变换时两边取自然对数，得：

$$\ln y = \ln d + \frac{b}{x} \quad (1-14)$$

$$\text{令 } y' = \ln y, a = \ln d, x' = \frac{1}{x},$$

$$\text{则有 } y' = a + bx' \quad (1-15)$$

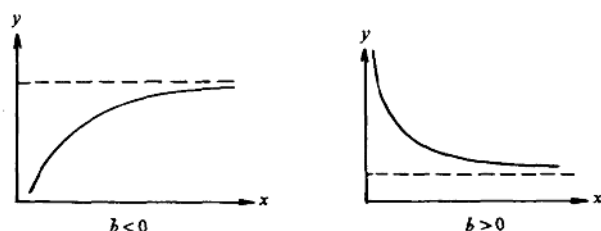


图 1-7 $y = de^{\frac{a}{x}}$ 的图像

现以宁芜地区中段单元内闪长玢岩出露面积比和矿床值的关系为例, y 为单元内已探明铁矿价值的对数值, x_1 为单元内出露的闪长玢岩面积与单元内面积比值的对数值 (见表 1-4、图 1-8)。

由图 1-8 可知, y 与 x_1 呈非线性相关。拟合成趋势曲线后, 该曲线与 $y = e^{bx}$ ($b > 0$) 的曲线类似, 按数据变换原理, 令 $y' = \ln y, a = \ln d$, 则可将非线性相关变换成近线性相关, 用线性方程组来代替指数函数。

表 1-4 宁芜中段单元矿床值与闪长玢岩出露面积比的对数值表

y	2.99	3.70	1.90	2.22	3.80	1.72	3.00	2.16	1.39	5.06	3.07
x_1	1.38	1.56	-0.12	0.80	1.55	0.85	1.36	0.86	-0.24	1.88	1.38
$y' = \ln y$	1.10	1.30	0.64	0.80	1.34	0.54	1.10	0.77	0.33	1.62	1.12
y	5.84	2.03	1.73	2.73	2.21	1.49	1.60	1.84	1.43	1.02	0.85
x_1	1.88	0.85	0.93	1.04	1.04	-0.12	0.18	0.76	0.18	-0.12	-0.82
$y' = \ln y$	1.76	0.71	0.55	1.00	0.79	0.40	0.47	0.61	0.36	0.02	-0.16

在拟合趋势曲线时, 如果在散点图上难以确定曲线的趋势, 那么可先用几种不同类型的函数作变换, 然后求出对应于各种变换后的相关系数, 选择其中相应相关系数 (绝对值) 最大的函数作变换。

以上列举的几种变换方法是最普通的方法, 但对原始数据模型通过其数据的变换后, 能形成方法数据模型满足预测方法模型相应的数学方法的要求并不多, 在这种情况下, 成矿预测人员要慎重考虑变换的方法和造成这种现象的原因。

其它方法如均匀化变换、角变换、平方根变换等在此不再叙述, 可参考有关文献。

成矿预测的实施在很大程度上是使用数据来实施的, 没有数据, 就不可能有成矿预测。在当前开展成矿预测工作的情况下, 地质数据显得越来越重要, 对地质数据处理和转换的方法研究工作, 受到更多的研究人员的重视。

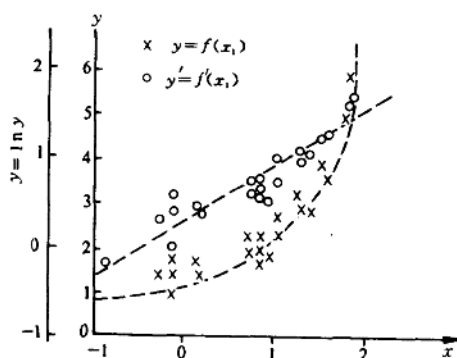


图 1—8 宁芜中段地区单元矿床值与闪长玢岩出露面积比散点图及变换
(据赵鹏大等, 1981)

第三节 单 元

成矿预测是通过矿床与地质条件之间的关系分析(数学的、经验的和感觉的)实施的, 又借用地质单元这一形式将矿床与地质数据观察值建立预测模型, 得以将已知矿床、潜在矿床及其相应的资源量作出预测评价。在这里, 单元是至关重要的。

单元只是为了统一观察值和取值范围, 将预测研究的范围(或是控制区、预测区)划分成矿等面积(或等体积)或面积相近的小区, 称之为单元。它是地质数据的取值单位。近年来, 随着我国成矿预测经验的积累, 对单元的认识越来越透彻, 解决了不少理论上的问题和实际中的技术问题。在一定的成矿预测范围内, 单元面积越大, 不含矿或含一个矿床的单元数越少, 含矿单元中与成矿有关的地质数据特征也发生相应的变化, 所以需要确定大小合理的单元进行预测。按我国当前的实际情况, 单元可分为:

1. 网格单元, 可以是正方形的, 也可以是长方形的;
2. 不规则单元, 也称地质体单元;
3. 穿透单元, 也称块段单元;
4. 立体单元;
5. 无单元型。

无单元型仅对一些特殊方法才使用, 如趋势面分析; 立体单元在立体预测工作中使用; 穿透单元对特殊产状形态的矿体才使用, 例如产在向斜中的矿体仅在两翼有露头时可使用这种形式; 当前大量地使用网格单元和不规则单元两类。网格单元划分简单, 数据取值和资料整理都比较容易, 但是它对自然地质体作了人为的割裂, 致使对某些地质现象的描述失去真实性。不规则单元克服了网格化单元固有的弱点, 但每个单元的大小都难以控制。通过对近年的实际经验总结, 不规则单元的基本优点有如下几方面:

1. 将完整的地质体作为研究对象, 可以较为直观地建立矿床与地质条件之间的关系;