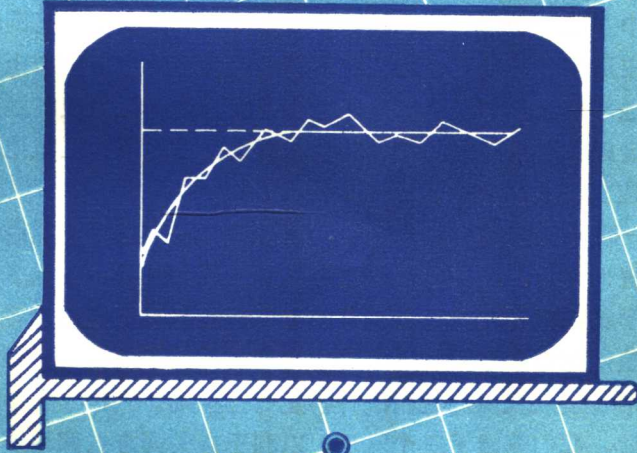


地质科技资料选编 (六十一)

地质统计学



地质部情报研究所

一九八〇年十月

编 者 的 话

地质统计学是利用统计学的某些原理和方法,探讨和解决地质学领域中的某些实际问题的一门边缘地质学科。它的主要任务是解决矿床从普查勘探到矿山开采过程中的各种储量计算问题。经过二十多年的发展,这门新兴学科已经建立了自己的理论基础和比较系统的方法体系,在解决实际问题、特别是矿石储量计算方面取得了明显效果。目前,国外一些工业和科学发达的国家都在继续大力研究,进一步发展它的理论和方法,扩大其应用范围。在我国,地质统计学的研究开始不久,亟待深化和提高。因此,我们编辑了这本译文选集,献给我国地质统计学工作者和广大地质技术人员。

集中收集的二十五篇论文是从五百余篇地质统计学文献中筛选出来的,大体分成两组:一组主要讲理论和方法,另一组着重谈实际应用。在理论和方法的介绍中,除了选取一些通俗易懂的文章供初学者参考外,重点放在转换函数、析取克里金和矿体模拟上。这些方法是法国数学形态中心在七十年代取得的重要科研成果,反映了当前国际上的研究水平。在地质统计学的实际应用部分,着重介绍了在不同成因类型的金属矿床上进行储量计算的实例,也反映了应用地质统计学方法选择最佳勘探方案和估计区域成矿远景的研究成果。由于篇幅所限,关于非金属和燃料矿产的储量计算,以及区域地质、区域地球物理和地球化学、地质制图等方面的应用,未作介绍。

地质统计学是一项十分实用的技术方法,必须遵循一些必要的原则才能取得良好效果。例如,地质与统计密切结合,对具体问题进行具体分析,就是一条重要原则。另一方面,它毕竟是一门轻年的学科,因而在理论上和方法上目前还都不可避免地存在一些缺陷和问题有待充实和完善。这些问题在本专辑所选的某些文章中也做了简略的介绍和讨论。了解并探讨这些原则和问题,对于发展我国的地质统计学是有裨益的。

在专辑正文之后,我们编录了三个附录,以利于有关同志查阅地质统计学外文文献。

在编辑这本集子的同时,在我所出版的《国外地质科技》杂志1980年各期上以“地质统计学讲座”为题发表了八篇译文,较通俗地对地质统计学做了系统介绍,适宜于初学者参考,本选集不再收录。

在本专辑编辑过程中,承蒙中国地质科学院矿床所四室数学地质组有关同志的大力支持和帮助,李裕伟、余金生同志从技术上审阅了全部译文,在此深表谢意。

由于编者专业和外文水平有限,这本集子错谬之处在所难免,务请读者批评指正。

编 者 谨 识

一九八〇年六月

目 录

地质统计学的基本问题	李裕伟 (1)
地质统计学及其在矿业上的应用	A. 马雷夏尔 (19)
正态和对数正态估计	J. M. 伦杜 (28)
矿石储量估计的泛克里金——基本理论及其在纳凡矿床上的应用	M. 达格伯特 M. 戴维 (40)
数学模型、样品集和矿石储量估计	A. G. 罗伊尔 M. J. 牛顿 (56)
析取克里金是条件期望的简单替代方法	G. 马特龙 (67)
预测块段品位分布: 转换函数	G. 马特龙 (76)
转换函数的实践: 数值法及其应用	A. 马雷夏尔 (85)
矿体条件模拟的地质统计学	A. G. 朱尼尔 (98)
矿石品位分布和条件模拟——地质统计学的两种方法	A. G. 朱尼尔 (112)
用地质统计学方法评价南非(阿扎尼亚)金矿山矿石的几个基本问题	D. G. 克里格 (116)
克里金的实践	M. 戴维 (125)
电子计算机化的地质统计学块段品位估计	B. 多德 (136)
借助数字图象显示器进行矿山设计和矿石储量估计	P. 多德 (152)
矿石储量估计的地质统计学方法与传统方法的对比	H. P. 努森 Y. C. 金 E. 米勒 (158)
地质统计学和地质学相互关系的新问题	F. P. 阿格特伯格 (165)
矿石储量估计中条件概率分布的一个应用实例	H. M. 帕克 (175)
洛林矿床铁矿石储量计算方法	L. 布伯尼塞克 A. 哈斯 (184)
变差图及其在软铁矿矿石类型解释中的作用	I. 杜塔 S. V. L. N. 拉奥 (196)
以若干实例说明斑岩铜矿品位和矿石-废石吨数估计的实践	M. 戴维 M. 达格伯特 J. M. 贝利尔 (205)

法国铝土矿矿床的地质统计学研究·····	A. 马雷夏尔 J. P. 鲁利埃(215)
爱尔兰白银厂莫古尔矿山回采工作面的矿石储量计算·····	M. 菲利翁(231)
南非(阿扎尼亚)金矿山采样间距的最优化·····	J. M. 伦杜(241)
法国地质矿产调查局用地质统计学确定矿体勘探的最佳方案·····	B. 莱利曼特(251)
应用矿化变化性的地质统计学特征选择稀有金属矿床的最佳勘探网度 ·····	Φ. P. 波利斯科依 Б. H. 巴尔尼科夫(255)

附录

附录 I 克里金和地质统计学现有英文文献述评·····	G. D. 贝尔 M. 里夫斯(262)
附录 II 地质统计学主要法文文献索引·····	编者(283)
附录 III 中英法地质统计学常见术语对照表·····	编者(286)

地质统计学的基本问题

李 裕 伟

引 言

六十年代初期法国数学家 G. 马特龙创立并发展了一门新的边缘地质学科——地质统计学，它的主要任务是解决从矿床勘探到矿山开采过程中的各种储量计算问题。地质统计学的核心是“克里金”，它是一种无偏的、误差最小的最优储量计算方法，这种计算方法是由南非采矿工程师 D.G. 克里格于 1951 年首次提出来的，故得名为“克里金”。简言之，克里金就是利用邻近若干个钻孔（或坑道）的样品品位来估计处于这些样品中间的某个块段（甚至某个点）的品位。使用这个方法，可以根据少量样品的品位资料把一个矿床中成千上万个开采块段的品位和储量统统计算出来。

目前，地质统计学已经形成一套完整的理论和方法体系，积累了大量的实际应用经验，受到西方国家许多采矿公司、大学和研究部门的重视。在以 G. 马特龙为首的法国巴黎矿业学院数学形态研究中心已经形成了一个强有力的法国学派。法国矿业界包括它在各地经营的矿山已把地质统计学作为储量计算的标准方法。美国、加拿大、澳大利亚、英国和一些拉美国家也在大力开展地质统计学的研究和应用。据报道，到 1975 年全世界已有 200 个矿山用地质统计学作了研究并发表了研究报告。一些计算机公司已着手建立以地质统计学方法为核心的储量计算软件。总的看来，地质统计学目前已成矿床勘探人员、矿山设计和开采人员不得不学习的一个新课题，应用地质统计学这一先进的储量计算方法已经处于势在必行的地步。

1977 年夏天，美国弗卢尔矿业金属有限公司 H. M. 帕克来华介绍了地质统计学，引起了我国一些矿床勘探、矿山设计和开采部门的注意。三年来，我国地质、冶金、二机部等部门有不少单位陆续开展了地质统计学研究，并取得了一些初步进展，但迄今为止尚未形成真正解决实际生产问题的能力。为了进一步推动我国地质统计学的研究，并把研究成果在生产实践中加以推广，有必要系统地介绍一下地质统计学的基本内容、它的工作方法和程序、在各类矿床中的应用，以及今后的发展趋势。

地质统计学的产生和发展

问题的提出

勘探地质学家的基本任务是，根据具体矿床的地质及其他特征布署勘探工程，取得有一定精度保证的储量。就勘探的意义而言，如果我们把矿床地质特征理解为矿床的变化性，把布署勘探工程理解为选择合适的勘探方法（工程类型、工程布置方向、工程间距），把储量

精度理解为储量误差，那么在整个地质勘探阶段中，勘探地质学家就需要不断地研究和处理矿床变化性、勘探方法和储量误差三者之间的关系。

矿床变化性、勘探方法和储量误差三者之间的关系，在储量规范中体现为勘探类型—勘探网度—储量级别的关系。多年来，我们就是通过执行储量规范调整三者之间的关系的。但是在执行规范过程中发现：

(1) 规范中在勘探类型、勘探网度和储量级别之间缺少严格的定量关系，由于这一缺陷，造成勘探人员、矿山设计和开采人员在使用规范对最终勘探成果进行审查和交接时往往争议不休；

(2) 地质人员长期没有认识到一个极为重要的统计原则：对储量误差的要求必须建立在储量计算块段大小的基础上，就是说，你必须指出是在多大的空间范围内谈论储量误差。例如，对同一勘探网度而言，在整个矿床范围内总储量误差很小，但对于该矿床的一个盘区来说储量误差则要大些，而对于组成盘区的块段来说，储量误差则更大。因此，如果不规定计算储量的空间范围的大小，我们就不可能谈论储量计算的精度。

由此可见，勘探工作实践中迫切需要解决的一个重要问题是，对于矿床中一定空间大小的区域（块段—盘区—矿床）寻找矿床变化性、勘探方法和储量误差三者之间的定量关系。

解决问题的途径

许多年来，地质勘探工作者试图找出一种新型的储量计算方法，以便反映矿床变化性、勘探网密度和储量误差三者之间的定量关系。

早在三十年代初期，苏联的勘探地质人员就开始使用数理统计方法来研究上述关系。他们发现，利用简单的区间估计公式，既可以计算理想的勘探网密度，也可以估计储量误差。但是他们随即认识到，要用简单的统计方法来解决复杂的地质问题是十分困难的，因为地质变量并不是纯粹的随机变量。

地质变量具有二重性——随机性和规律性。例如在一个矿床中，金属品位的变化在空间上总是具有这样或那样的规律性；但是，在这种规律变化的背景上又迭加了一些偶然的、随机变化的因素，使我们不能确切地指出某一空间位置的精确品位值。由于这种随机变化和规律变化的迭加，使我们不能使用经典的统计方法来解决地质勘探问题。苏联的勘探地质学家从三十年代到六十年代曾作了很大努力来分离随机变化和规律变化，产生了类似于变差图的样品空间相关的思想，甚至指出地质变量是随机函数而不是随机变量，但并没有取得在勘探地质学中成功应用统计方法的系统成果。

我们今天看到，在西方发展起来的地质统计学才比较成功地解决了在地质变量二重性变化条件下使用统计方法的问题。这首先要归功于四十年代末变差图的出现，变差图描述了地质变量的随机变化和空间规律变化，这就为在矿床勘探中使用统计方法铺平了道路。此外，几乎在这同一时期，D.G.克里格根据南非金矿山的勘探和开采工作实践得出了一种认识：在一个矿床或矿体中，样品或块段之间存在着某种几何关系。因此，他提出了利用相邻许多块段的平均品位估计中心块段的简单回归模型。克里格的这种原始回归模型，现在已发展成为一

套相当严密的地质变量空间插值技术——克里金，利用它可计算一个矿床任意具体位置的储量和品位。克里金加上变差图构成了地质统计学萌发的“胚胎”。

到了五十年代末，G. 马特龙研究了 D. 克里格 H.S. 西奇尔等人在南非的工作成果，在用随机函数这一概念描述随机变量的空间变化的基础上，提出了区域化变量的概念。马特龙把克里格等人的研究成果进一步理论化和系统化，从而形成了真正的地质统计学。

概而言之，从勘探角度来说，地质统计学就是在地质变量具有二重变化的条件下建立的一套解决地质问题的统计方法。它把矿床或矿体中的地质特征看成是具有二重变化的区域化变量；利用变差图来描述区域化变量的随机变化和规律变化，然后根据变差图所提供的矿床变化性进行克里金插值，计算出矿床中所有块段的品位和储量。克里金是一种无偏的、误差最小的最优化储量计算方法，在计算储量的同时我们可以获得一个相应的误差估计。区域化变量、变差图和克里金构成了地质统计学的三大支柱。利用地质统计学方法就可能在地质变量二重性的条件下正式地、系统地研究矿床变化性、勘探网度和储量误差三者之间的定量关系。

现在，地质统计学尚在迅速发展，如果说过去建立的传统地质统计学是研究某个变量的线性特征的话，那么今天地质统计学已经发展到了研究复杂组合的非线性问题，从而产生了转换函数、条件模拟等许多较为复杂的方法。

地 质 统 计 学 的 优 点

地质统计学储量计算方法具有下列明显的优点：

(1) 在许多情况下，所计算的品位和储量数字比传统的方法要精确得多。例如，某已采完的铜矿山，根据勘探资料用地质统计学方法预测的铜平均品位为 2.14%，经营该矿山的公司使用传统方法预测的铜品位为 2.51—3.20%，而根据出售的铜计算出的实际品位为 2.04%。再如，美国亚利桑那州有一个矿山，用传统方法的一个模型预测有 1600 万吨物料，其中 82% 是矿石，18% 是废石，用地质统计学方法预测则分别为 60% 和 40%，最后按开采资料算得实际比例为 59.4% 和 40.6%。这些例子说明，用地质统计学计算能达到相当精确的程度。

(2) 传统的储量计算方法对勘探资料的利用是不充分的，而克里金则可最大限度地利用勘探工程所提供的信息。克里金在计算矿床中某块段的储量时，不仅考虑了落在这个块段中的样品，而且还利用了许多落在块段外的邻近样品。这正是克里金能够提高储量计算精度的一个重要的原因。

(3) 传统方法提供的是若干个勘探块段的储量，这种块段对采矿设计而言范围太大，而且很不规则，设计部门难以使用。用地质统计学方法可以分别计算出矿床中所有最小开采块段的储量和品位（在基本确定开采方法的条件下）。这样就使勘探报告所提供的储量能更好地满足矿山设计的要求，从而有利于矿床勘探、矿山设计和矿山开采三个阶段之间的衔接。

(4) 地质统计学方法建立了矿床地质特征、勘探方案和储量精度三者之间严格的数学关系，因此，我们可以用它来研究一个矿床的合理勘探方案，如工程类型（钻孔、坑道、槽井探等）、工程摆布方向、勘探网密度、取样方法等。特别是在一个勘探阶段向另一个勘探阶段过渡时，地质统计学可以预测达到某一精度所需要的工程量，也可以预测勘探储量估计的

精度。

由于从勘探开始直到矿山开采完毕的整个过程中要不断地进行储量计算，所以用作储量计算的地质统计学方法也贯穿这一过程的始终。在勘探期间，利用它可寻求最合理的勘探方案和计算勘探储量；在矿山设计过程中，可利用它来探讨边界品位、矿床储量、开采方案、利润和资源利用之间的关系，以选择合理的边界品位和开采方法；在采矿期间，可利用它来作矿石数量和质量的短期预测，以便合理地安排矿山生产。

品 位 的 概 率 分 布

在地质统计学中，地质变量，尤其是品位的概率分布是一个既有理论意义又有实用意义的概念。所谓品位的概率就是指一个样品可能出现某一个随机选择的值的可能性。人们研究过许多种代表品位分布的模型，其中最有意义的是正态分布和对数正态分布，前者构成了几乎现有一切统计方法的基础，后者也很早为地质学家为所注意。但是在许多情况下，矿石品位既不是正态分布，也不是对数正态分布，往往成为一种形态不自然的偏斜，甚至多峰的分布。这多半是由于子样混杂有多个统计总体所致。多个总体可以理解为是多种地质环境（或过程）的产物。例如，许多矿床由不同的矿石类型组成，每一种矿石类型很可能具有各自的正态或对数正态分布，但混杂在一起，互相干扰和歪曲，从而形成了上述的混合分布。

前面只是抽象地讨论了品位分布，但我们要研究的是实际品位分布，即样品的品位分布、不同大小的块段的品位分布、包括许多块段的盘区的品位分布，以及整个矿体或矿床的品位分布。在讨论这种实际的品位分布时，我们必须注意样品或块段的品位的真正涵义。

在具体的地质、采矿活动中，样品或块段都具有一定的体积，因此，一个样品实际上是由许多细小的点组成，所谓样品品位就是该样品中所有这些点品位的平均值。同样，一个块段的品位可以看成是该段块所占据的体积内所有点品位的平均值（当然也可以看成是所有样品品位的平均值）。由此可见，所谓品位的概念，就是在一定大小的体积单位中点品位的平均值的概念，在这里确定体积是最重要的。这样，研究块段品位的分布就成了研究块段点品位平均值的分布。

与此相联系的是“支集”的概念。如上所述，地质统计学研究的对象不是一个理论的点而是一个实体，这个实体具有许多影响随机变量取值的现实特征，我们把它们称为对象的支集；其中那些描述对象的大小和形态的现实特征称为对象的几何支集。因此，我们在讨论对象的品位分布的时候总是指在一定几何支集意义下的品位分布，如一定尺寸和形态的样品品位的分布，一定尺寸和形态的块段品位的分布，等等。正如上面所指出的，这里所说的品位分布实际上都是点品位平均值的分布。

通常，如果样品的体积相对于块段的体积来说可以忽略不计的话，为了简化计算，我们也不妨把样品看作一个“点”，于是可以以样品为起点计算块段的品位和方差，从而解决由样品品位的概率分布讨论块段品位的概率分布的问题。当然要做到这一点还要解决其他许多问题。

品位的概率分布无论在估计矿床的总储量还是在估计块段的储量方面，都是一个强有力的工具。关于这一点，本专辑中将有许多文章加以介绍，例如，用品位-吨位曲线估计矿床

的总储量，用对数正态分布估计一个块段中矿石和废石的比例，等等。

研究对象—区域化变量

在讨论区域化变量之前，我们先弄清下列两个概念：

随机变量：它是一组可能的数值，其中与每个数值相对应都有一个概率。每次观测（或取样）的结果是一个确定的数值，叫做随机变量的实现。

随机函数：它是一组可能的函数，其中与每个函数相对应均有一个概率。每次观测（或取样）的结果是一确定的函数 $z(x)$ ，我们称之为随机函数的一个实现。当随机函数的座标 X 固定时，各随机函数的可能取值即构成一个随机变量。

区域化变量就是在某个区域中定义在一个确定的空间位置上的随机变量。不同空间位置的区域化变量之间呈现某种空间规律变化，因此，这些随机变量是空间相关的，而这种相关正反映了矿化过程的连续性和关联性。在储量计算中，如果矿体算是区域，那么矿石品位、厚度、密度、线储量或其他任何矿石特征都可以是区域化变量。马特龙把区域化变量定义为一个三维空间中的座标函数，这个函数具有规律性的空间变化（结构变化）和随机变化的两重性质。区域化变量不能用一个常规的函数来确定地解析，而只能用一组变差图来同时描述它的规律性的空间变化和随机变化。

区域化变量被看成是一个随机函数 $Z(x)$ 。从随机函数的观点，对区域化变量提出了三种假设，由之导出了相应的地质统计学方法。

弱平稳假设

在整个研究区域内，区域化变量 $Z(x)$ 的期望是相等的，协方差将只取决于滞后，这时我们可用以下常规的形式定义其数学期望、协方差和方差：

$$E[Z(x)] = m$$

$$\text{Cov}(x, x+h) = E\{[Z(x) - m][Z(x+h) - m]\} = K(h)$$

$$\text{Var}[Z(x)] = E\{[Z(x) - m]^2\} = K(0)$$

上述公式中： m 为样品品位， h 是二维空间的一个向量，其长度为 h 。

内蕴假设

实际上，许多自然现象多半很难满足弱平稳假设。但这时若不考虑函数 $Z(x)$ 自身而考虑其增量 $Z(x) - Z(x+h)$ ，即该增量可能满足平稳假设，由此可认为随机函数 $Z(x)$ 是内蕴平稳的。这时定义增量的数字期望和方差为：

$$E[Z(x+h) - Z(x)] = m(h)$$

$$\text{Var}[Z(x+h) - Z(x)] = 2\gamma(h)$$

(1)

式中 $\gamma(h)$ 为半方差，在下一节变差图中将作详细介绍。

内蕴假设的意义在于，在该假设下，我们得不到 $Z(x)$ 的协方差，但却可以计算出一个半方差 $\gamma(h)$ 。在进行储量估计时利用的是 $\gamma(h)$ ，从而避免使用协方差 $K(h)$ 。正如马特龙所指出的，一开始就使用半方差而不使用协方差是一件十分幸运的事，它使地质统计学家在平

稳性和遍历性问题面前没有陷入绝境。

但是，如果随机函数 $Z(x)$ 是二阶平稳的（它同时当然也是内蕴平稳的），则可以据弱平稳假设计算出协方差 $K(h)$ ，这时半方差可直接由协方差导出：

$$\gamma(h) = K(0) - K(h) \quad (2)$$

泛克里金假设

如果在一个限定范围的邻域 U 内，区域化变量 $Z(x)$ 或其增量 $Z(x+h) - Z(x)$ 的二阶矩具有某种平稳性质，但其数学期望在该邻域中却不是平稳的，而具有一定的规律变化，则这时在邻域 U 中定义区域化变量的数学期望和方差为：

$$E[Z(x)] = m(x) \quad (x \in U)$$

$$E[Z(x) - Z(y)] = m(x) - m(y) \quad (y = x+h, x, y \in U)$$

$$m(x) = \sum_{i=0}^k a_i f_i(x) \quad (3)$$

$m(x)$ 为域 U 中任一点的数学期望，即漂移； $f_i(x)$ 是根据任一点的三维空间坐标结合成的一个趋势项； a_i 为每一个趋势项的系数。

$$\text{Cov}[Z(x), Z(x+h)] = K(h)$$

或

$$\text{Var}[Z(x+h) - Z(x)] = 2\gamma(h)$$

上面是目前在地质统计学中对区域化变量作出的三种最普遍的假设。使用每种假设所提供的参数估计方法，可以计算克里金方程和误差方程所需的基本参数，从而实现了对矿床或块段储量的克里金估计。

在实际的地质现象中，弱平稳假设是难得见到的，泛克里金假设处理起来非常繁琐，故目前最普遍使用的是内蕴假设意义下的克里金技术。但是，通过地质人员和数学家的共同努力，往往可以把一个泛克里金假设的地质现象化为一个能满足内蕴假设的地质现象，从而大大简化了计算过程。

基本工具 — 变差图

地质统计学工作过程的第一步就是作出变差图。变差图既描述了地质变量的空间规律变化，又描述了迭加在这种规律变化之上的随机变化，因此，它显然是一个迄今为止描述矿床变化性的最理想的数学工具。

在经典统计学中，我们用方差

$$\text{Var}(Z) = \frac{1}{N-1} \sum_{i=1}^N (Z_i - \bar{Z})^2$$

来描述随机变量的变化性，这时仅存在随机的变化，而不存在任何空间的、规律的或结构的变化。在地质统计学中，我们实际上仍然要用方差的概念来描述地质变量的变化性，但这时计算方差时将要考虑到样品空间相关的牵制。特别是在内蕴假设的条件下，我们假定两个相邻的、距离为 h 的样品品位之差是平稳的，于是这个品位差便构成了我们计算方差的基础，其计算公式与上式在形式上雷同，

$$\gamma(h) = \frac{1}{2(N-h)} \sum_{i=1}^N [Z(x_i) - Z(x_i+h)]^2 \quad (4)$$

对一个确定的 h , $\gamma(h)$ 是一个值, 称为半方差。对于一系列滞后 h , 据上式可算出一系列相应的值。在 $\gamma(h)$ 对 h 的坐标图上, 每个值都相当一个点, 将各点联结得经验半变差曲线 (图 1)。

在变差图建立之后, 需要通过研究各个方向上的变差曲线来了解不同方向上矿床的变化特征。这时可能出现各向同性和各向异性两种情况。如果是各向同性, 则各方向上变差曲线在形态上和位置上均很接近, 这时问题比较简单, 将各方向上的变差曲线予以简单平均, 即可作为一个代表矿床总变化的变差图予以使用; 但当出现各向异性时, 问题就比较复杂, 需要设计一些方法来合成各方向的变差曲线。

得到经验变差图后, 必须把它和某种理论模型相拟合。理论模型可以分为二组: 一组是理论上的 $\gamma(h)$ 随着距离 h 的增大而增大; 另一组是 $\gamma(h)$ 起初增大, 而后趋向于在一个常数 γ 值上水平地延伸。后者被认为有一个基台, 它的值通常用 $C + C_0$ 表示,

没有基台的模型中较重要的是戴维杰斯模型:

$$\gamma(h) = 3a \log_e h \quad (5)$$

这种模型是对数线性的, 因而简便易行。

有基台的模型中最重要的是球形模型 (图 1):

$$\begin{aligned} \gamma(h) &= C \left[\frac{3h}{2a} - \frac{h^3}{2a^3} \right] + C_0, & h \leq a \\ \gamma(h) &= C + C_0, & h > a \end{aligned} \quad (6)$$

式中 a 、 C 、 C_0 是球形模型的几个基本参数。 a —变程, 代表样品的最大影响范围, 超出这个距离, 两样品间不再存在相关关系。

C —结构随机变化的极大值。所谓结构随机变化, 指的是取决于两点之间距离的随机变化, 当两点间的距离超过变程 a 时, 则结构随机变化恒为 C 不变。

C_0 —块金效应。这是一种非结构随机变化, 同两点间距离无关。人们早就注意到, 对许多变差图而言, 当两点间的距离为 0 时, 半方差不为 0, 而有一个稳定的 C_0 值。块金效应既可能是由于取样、加工和化验的误差所造成, 也可能是由于矿化不连续性引起的瞬变而产生。 $C + C_0$ —基台, 等于总方差。如果把区域化变量的随机变化分解为两种成分, 一种是随距离而改变的结构随机变化, C 是这一变化的极大值; 另一种是不随距离而变化的非结构随机变化

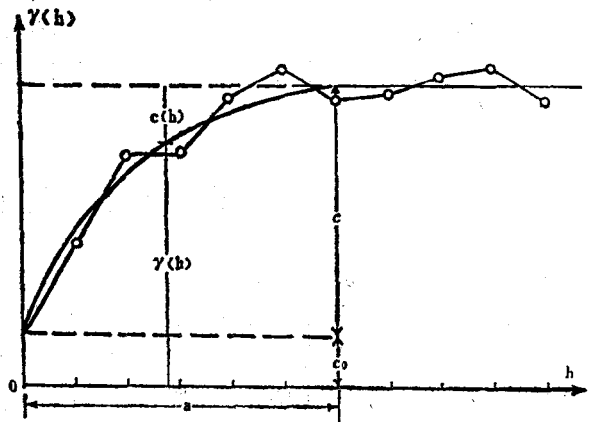


图 1 经验变差图及相应的球形模型

(C_0), 那么, 达到变程 a 时两种成分随机变化之和达到极大值 $C + C_0$, 而规律性变化为零。

$C(h)$ —滞后为 h 时的结构变化，即样品之间的相关变化。

下表列出变差图的几种理论模型。

没有基台的模型		
线性模型	$\gamma(h) = \rho h$	
广义的线性模型	$\gamma(h) = \rho h^\lambda$	$0 < \lambda < 2$
戴维杰斯模型	$\gamma(h) = 3a \log_e h$	
有基台的模型		
指数模型	$\gamma(h) = C[1 - \exp(-h/a)]$	
球形模型	$\gamma(h) = C \left[\frac{3h}{2a} - \frac{h^3}{2a^3} \right]$	$h < a$
	$= C$	$h > a$
块金效应	$\gamma(h) = 0$	$h = 0$
	$= C_\sigma$	$h > 0$

变差图所表达的基本上还是一个方差的概念，但它不是一种常规的方差，而是在地质统计学意义下的方差——变方差。也就是说，变差图所描述的方差是随着距离、方向和支集的变化而变化的。

对地质统计学而言，变差图几乎集中了矿床变化的全部信息，它提供了实施各种地质统计学估计的有力基础。我们在估计储量和相应的误差、模拟品位空间分布的时候，都要向变差图索取依据。正因为如此，在建立变差图时要特别谨慎。如果变差图一旦出了重大错误，就会影响到整个储量计算。

变差图的实际建立，主要不是数学问题而是地质问题，这是因为：（1）经验半方差的计算总是非常简单，地质人员能很快学会；（2）只有地质人员才能解释变差图，如果变差图不加解释就投入使用，那是危险的；（3）地质人员能在方向选择、钻孔搜索、样品组合等计算变差图的基本问题上出谋划策，而且当使用了不当的工程而导致错误或出现不正常的情况的时候，只有他们能及时发现，追查原因，并予以纠正。鉴于以上原因，地质人员掌握变差图具有十分重要的意义。

基本的方差构成模型

虽然变差图从本质上说是一个方差的概念，但我们却往往不能用它来描述实际的方差，因为变差图一般是在零支集（点）的意义下作出的，而我们要计算的方差总是涉及到一个块段（或样品）这样一些具有一定大小和形态的几何单位；而且，当利用不等权的样品品位来推断块段品位的时候，点方差同块段（或样品）方差之间的关系就更为复杂。

但是，在地质统计学中，变差图毕竟提供了构造实际方差的基本元素。利用这些元素可以形成在各种支集下的样品或块段的品位的方差。当然，当一个样品的体积被忽略，可以视为一个点的情况下，从变差图上读取的值即是实际的方差。

下面是几个重要的方差构成模型：

1. 离散方差 又称块段内方差，指的是一定体积的块段（或样品）内点品位的离散。

在体积为 V 的块段 A 中定义一个向量 $|x-y|$ ，其中 x 和 y 分别为该向量的两个端点。我们连续地改变端点位置，对每次改变均从变差图获得一半方差值 $\gamma(x-y)$ ，对该变差函数进行体积积分并取平均值，则得块段 A 内的离散方差为：

$$\sigma^2(O/V) = \frac{1}{V^2} \int_V dx \int_V \gamma|x-y| dy = \bar{\gamma}(V) \quad (7)$$

这个公式又称为 F 函数，以 $F(y)$ 表示。这是一个十分重要的基本的方差构成公式，据此可以利用变差函数计算的半方差值 $\gamma|x-y|$ 构成任意块段的离散方差。而任意块段的离散方差正是计算大块段中以小块段为单位的离散方差的基础，其公式为：

$$\begin{aligned} \sigma^2(v/V) &= \sigma^2(O/V) - \sigma^2(O/v) = F(V) - F(v) \\ &= \bar{\gamma}(V) - \bar{\gamma}(v) \end{aligned} \quad (8)$$

式 (8) 是非常有用的，它说明了块段大小同方差变化的关系。

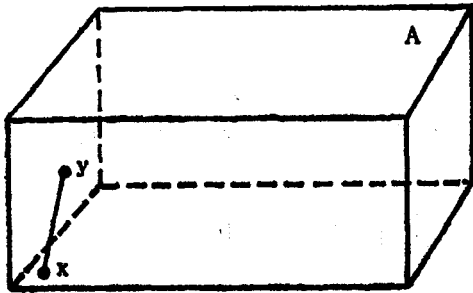


图 2 由向量 $|x-y|$ 扫描整个块段形成离散方差

2. 两块段的离散协方差 指的是两块段（或块段与样品）间点品位的离散。

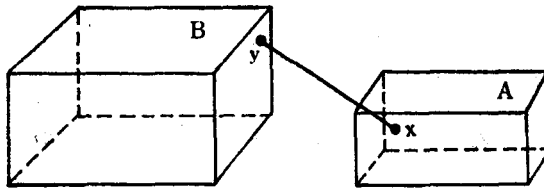


图 3 由两端点 x, y 分别在块段 A, B 中的向量 $|x-y|$ 扫描体积 A, B 形成离散协方差

设有块段 A 和 B （或样品 A 和 B ），或样品 A 和块段 B ，其体积分别为 V 和 v ，定义一个向量 $|x-y|$ ，其中 x 和 y 分别为该向量的两端点，一点在 A 中，一点在 B 中。类似于公式 (7)，我们得一个 6 重积分：

$$\sigma^2(O/V, O/v) = \frac{1}{vV} \int_V dx \int_v \gamma|x-y| dy = \bar{\gamma}(V, v) \quad (9)$$

3. 延伸方差 如果我们要根据一个小体积的样品品位 $Z(v)$ 去估计一个大体积的块段的品位 $Z(V)$ ，则将出现一个估计方差 $\text{Var}[Z(V) - Z(v)]$ ，这就是所谓延伸方差。延伸方差描述了一个样品离开它自身的体积向外部扩展而产生的推断方差。

据式 (7) 和 (9)，首先计算出两个体积的三种离散方差：

$$\sigma^2(O/V) = \bar{\gamma}(V) = \frac{1}{V^2} \int_V dx \int_V \gamma|x-y| dy$$

$$\sigma^2(O/v) = \bar{\gamma}(v) = \frac{1}{v^2} \int_v dx \int_v \gamma |x-y| dy$$

$$\sigma(O/O, V/v) = \bar{\gamma}(V, v) = \frac{1}{vV} \int_v dx \int_v \gamma |x-y| dy$$

据公式(8), 以体积 V, v 为单位的品位方差及其协方差为:

$$\sigma^2(V/D) = \bar{\gamma}(D) - \bar{\gamma}(V)$$

$$\sigma^2(v/D) = \bar{\gamma}(D) - \bar{\gamma}(v)$$

$$\sigma(V, v/D) = \bar{\gamma}(D) - \bar{\gamma}(V, v)$$

则延伸方差为:

$$\begin{aligned} \text{Var.}[Z(V) - Z(v)] &= \text{Var}[Z(V)] + \text{Var}[Z(v)] - 2\text{Cov}[Z(V), Z(v)] \\ &= [\bar{\gamma}(D) - \bar{\gamma}(V)] + [\bar{\gamma}(D) - \bar{\gamma}(v)] - 2[\bar{\gamma}(D) - \bar{\gamma}(V, v)] \\ &= 2\bar{\gamma}(V, v) - \bar{\gamma}(V) - \bar{\gamma}(v) \\ &= 2 \frac{1}{Vv} \int_v dx \int_v \gamma |x-y| dy - \frac{1}{V} \int_v dx \int_v \gamma |x-y| dy \\ &\quad - \frac{1}{v^2} \int_v dx \int_v \gamma |x-y| dy \end{aligned} \quad (10)$$

该公式是计算块段推断误差的基础。但是, 在实际的被估块段中, 所有的已知信息并不是集中在一个样品上, 而通常是用散布在块段内外的若干个样品来估计该块段的品位, 这时为简便起见, 我们把每个样品看成一个点, 被估块段仍被视为一个连续的体积, 则根据若干个样品估计块段品位的延伸方差, 公式为:

$$\begin{aligned} \sigma_{sv}^2 &= \frac{2}{nV} \sum_{i=1}^n \int_v \gamma |x_i - x| dx - \frac{1}{V^2} \int_v dx \int_v \gamma |x-y| dy \\ &\quad - \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \gamma |x_i - x_j| = 2\bar{\gamma}(S, V) - \bar{\gamma}(V) - \bar{\gamma}(S) \end{aligned} \quad (11)$$

式中 $\bar{\gamma}(S, V)$ 为样品同块段的离散协方差, $\bar{\gamma}(V)$ 为块段内部的离散方差, $\bar{\gamma}(S)$ 为样品的离散方差。这时各样品在计算块段品位时是等权的, 即块段品位等于各样品品位的算术平均值:

$$\bar{Z}(V) = Z(v) = \frac{1}{n} \sum_{i=1}^n Z(x_i)$$

4. 克里金方差 克里金方差是上述延伸方差的概念在样品不等权意义下的直接推广。

根据样品计算块段品位的加权平均公式为:

$$\hat{Z}(V) = \sum_{i=1}^n a_i Z(x_i)$$

则该随机变量线性组合的方差为带有加权系数的所有随机变量两两协方差之和:

$$\sigma_s^2 = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{Cov}[Z(x_i), Z(x_j)] = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma |x_i - x_j| \quad (12)$$

样品同块段的离散协方差为:

$$\sigma_{zV}^2 = \sum_{i=1}^n a_i \text{Cov}[Z(V), Z(x_i)] = \frac{1}{nV} \sum_{i=1}^n a_i \int_V \gamma(x_i, -x) dx \quad (13)$$

同式 (11) 相类似, 得出样品加权平均的块段推断品位方差的公式——克里金方差估计公式:

$$\begin{aligned} \text{Var}[\hat{Z}(V) - Z(V)] &= \frac{2}{nV} \sum_{i=1}^n a_i \int_V \gamma |x_i, -x| dx - \bar{\gamma}(V) - \\ &\quad - \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma |x_i, -x_j| = \sigma_{zV}^2 \end{aligned} \quad (14)$$

上述公式是在 $Z(x)$ 二阶平稳的假设下导出的。马特龙证明, 在内蕴假设下, 用半方差来合成克里金方差的公式 (14) 仍然是正确的, 这可以避免使用协方差。

在实际应用中, 由于涉及到大量积分计算, 这种用半方差 $\gamma |x-y|$ 来合成样品或块段方差的办法是非常繁琐的。但对一些常用的变差图模型 (如球形模型和戴维杰斯模型) 而言, 有些现成的图、表可供查用。

地质统计学的一个重要任务就是在样品相关的条件下, 研究实际方差的构成规律, 它包括: 方差的构成要素——变差图、体积-方差关系和由样品到块段的估计误差。上述这些对相关样品方差的深入研究奠定了克里金估计的坚实基础。

基本插值模型—克里金

对金属矿床而言, 储量计算的参数主要是品位。在勘探工作结束之后, 我们只获得若干个钻孔的矿石品位, 而矿床内部绝大部分空间内品位的值是未知的。储量计算的主要过程实际上是如何根据钻孔的品位来估计矿床的品位, 换句话说, 如何把占矿床极少部分的钻孔品位传递给整个矿床, 这实际上是一个插值问题。

地质统计学中的克里金方法是一种高技巧的统计插值模型, 它能够做到把一个矿床中有限的钻孔数据传递到矿床的任一局部地区。例如, 如果在矿床勘探时打了 100 个钻孔, 按矿山开采要求划分了 10000 个块段, 那么利用克里金模型能把这 100 个钻孔数据轻而易举地传递到 10000 个块段中去。换言之, 每个块段的品位和储量都能被估计出来。

克里金是在无偏性和最小估计方差两个条件保证下进行估计的。其实质在于, 将一个被估块段邻近的若干个样品的品位传递给该块段, 而不论这些样品是否在该块段之中。每个样品对该块段的影响由变差图所提供的距离-半方差关系来确定; 由于样品是小支集, 块段是大支集, 故这种传递也是以平均值的方式进行的, 即被估块段的品位是邻近若干个样品按平均的距离-半方差关系所确定的权加权的平均值。

克里金的基本模型是

$$\hat{Z}(v) = a_1 Z(x_1) + a_2 Z(x_2) + a_3 Z(x_3) + \dots + a_n Z(x_n) \quad (15)$$

式中 $\hat{Z}(v)$ 为块段的估计品位, $Z(x_i)$ 是为了估计块段 V 的储量所采用的样品品位, a_i 为克里金系数。

如果钻孔样品的体积不能忽略，则不可视为点样品，这时必须计算每个样品及每两个样品之间的平均半变差 $\bar{\gamma}_{ij}$ ，并用它来代替正规方程组中的 γ_{ij} 值。

如果在一个块段的范围内矿化存在着趋势，那么这种情况就被说成是不平稳的，这时就必须使用推广的克里金，即“泛克里金”。

泛克里金方程的推导同上述基本克里金（在全域满足内蕴假设）没有根本不同之处。本专辑“矿石储量估计的泛克里金—基本理论及其在纳凡矿床上的应用”一文对泛克里金作了详细讨论，这里不再赘述，仅列出最终的泛克里金正规方程组，其中第二个方程为泛克里金的无偏条件：

$$\begin{cases} \sum_{j=1}^n \gamma_{ij} a_j - \sum_{l=0}^k \mu_l f_l(x_i) = \bar{\gamma}_{i,v} \\ \sum_{i=1}^n f_i(x_i) a_i = \frac{1}{V} \int_V f_i(x) dx \end{cases} \quad (20)$$

泛克里金的数学表达式很简单，但实行起来有许多困难，目前还不是一个成熟的方法。在实际应用中，应尽可能把一个可能属于泛克里金的问题从地质上化为一个属于基本克里金的问题，或许在效果上和经济上更为合适。

迄今为止，我们仅讨论了用一个变量自身空间变化的信息来进行克里金估计。能否用另一个或多个不同的变量来实施对某个变量的克里金估计呢？这是一个很有实际意义的问题。例如，在矿床勘探阶段，仅对主要金属元素作了大量基本分析，但在若干年以后，当矿床投入开采时，发现矿石富含某种具有工业意义的伴生元素，这在矿床勘探史上是屡见不鲜的事。但这时岩心和样品已部分散失，重新化验已不可能，重新打钻又不合算。如果我们能根据矿床中主要元素的大量基本分析数据计算某伴生元素的储量或计算任意局部块段某伴生元素的储量，那将有巨大的经济意义。现在，地质统计学中发展起来的协克里金就可解决这个问题。

协克里金模型为

$$\hat{Z}^{(l)}(v) = \sum_{i=1}^n a_i z^{(k)}(x_i)$$

式中 $\hat{Z}^{(l)}(v)$ 为块段 v 中第 l 种元素的估计品位， $z^{(k)}(x_i)$ 为第 i 个邻区样品第 k 种元素的含量。这是一个用 n 个样品中第 k 种元素品位求块段 v 第 l 种元素品位的克里金模型。

求解克里金系数将用到互变差图，建立经验互变差图的半方差被定义为：

$$\gamma^{(kl)}(h) = \frac{1}{2N} \sum_{j=1}^N [z^{(k)}(x_i) - z^{(k)}(x_i + h)][z^{(l)}(x_i) - z^{(l)}(x_i + h)] \quad (22)$$

式中 $\gamma^{(kl)}(h)$ 是在滞后为 h 时变量 k 与 l 的半方差， N 为参与计算变差图的样品对数。

协克里金的估计方差为：

$$[\sigma_{s_v^{(l)}}]^2 = 2 \sum_{i=1}^n a_i \bar{\gamma}^{(kl)}(x_i, v) - \bar{\gamma}^{(l)}(v) - \sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma^{(kl)}(|x_i - x_j|) \quad (23)$$

式中 $[\sigma_{s_v^{(l)}}]^2$ 为被估块段第 l 种元素品位的估计方差， $\gamma^{(kl)}(x_i, v)$ 为第 i 个样品同被估块段关于元素 k, l 的平均互半方差， $\bar{\gamma}^{(l)}(v)$ 为被估块段中第 l 种元素的平均半方差， $\gamma^{(kl)}$