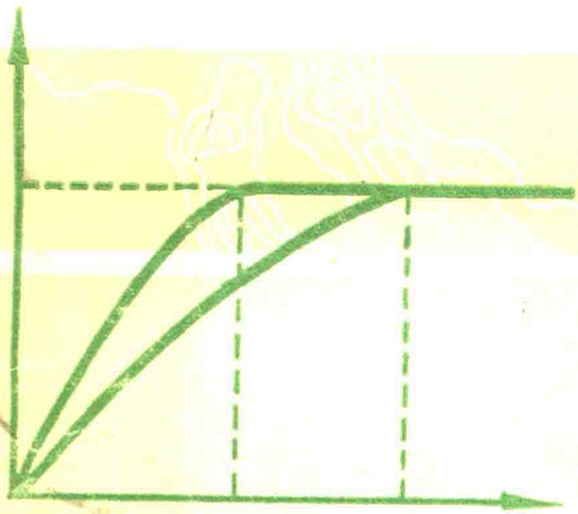
等学校教学用书

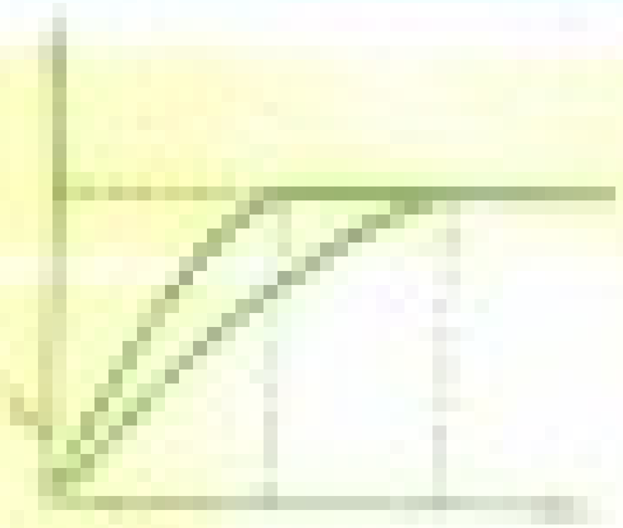
地质统计学及其应用



孙洪泉 编

中国矿业大学出版社

地质统计学及其应用



高等学校教学用书

地质统计学及其应用

孙洪泉　编

中国矿业大学出版社

内 容 提 要

本书系统介绍了线性地质统计学的基本理论和基本方法，详细阐述了普通克立格和泛克立格的方法和应用，并对非参数地质统计学的指示克立格法、非线性地质统计学的析取克立格法以及条件模拟等作了简介。书中引用了地质统计学在煤田地质勘探、储量计算等方面的研究成果作为应用实例，详细阐明了地质统计学的应用方法。书中还附有在长城0520微机上实现的地质统计学程序，主要包括：数据处理、屏幕绘图、计算实验变差函数、理论变差函数的拟合、块段克立格估值等程序。每个程序都有计算实例和详细的使用说明。

本书可作为高校地质或采矿类专业的本科生、研究生的教学用书，也可作为数学地质工作者的参考书。

责任编辑：瓮立平
责任校对：杜锦芝

高等学校教学用书
地质统计学及其应用
孙洪泉　编

中国矿业大学出版社出版
江苏省新华书店经销　中国科学院开封印刷厂印刷
开本787×1092毫米 1/16　印张18.25 字数433千字
1990年10月第一版　1990年10月第一次印刷
印数：1—2000册

ISBN 7-81021-380-6

P·13(课)　　定价：3.65元

前　言

地质统计学是六十年代初期法国著名学者G.马特隆（G.Matheron）教授创立并发展起来的一门新兴边缘学科，是数学地质的一个独立分支。地质统计学的核心是“克立格”。它是一种无偏的最小误差的储量计算方法。这种方法是由南非采矿工程师D.G.克立格(D.G.Krige)于1951年首次提出的，故命名为“克立格”法。

地质统计学虽然是从研究矿产储量计算以及误差估计问题产生和发展起来的，但它所提出的数学模型和方法同样适用于许多不同的学科领域。可以说凡是要研究空间分布数据的空间结构性和随机性、并对它进行最优线性无偏内插估计时，均可应用地质统计学。

为了使地质统计学在我国煤田地质勘探、矿井设计、煤矿开采及其他有关学科领域内更好地推广应用，编者在从事地质统计学的教学和科研的基础上编写了这本教材，供煤田地质勘探、采矿等有关专业开设地质统计学课程使用，也可作为从事地质统计学研究人员的参考书。为了使读者更好地应用地质统计学，书中还给出了地质统计学的计算机程序（主要有数据处理、变差函数拟合、交叉验证、克立格估值等）。

全书共有十章，主要介绍线性地质统计学的基本内容。第一章是绪论；第二章是线性代数、概率统计基础知识；第三章至第七章为线性地质统计学的基本理论、基本方法，主要是普通克立格法；第八章是泛克立格法；第九章是指示克立格、析取克立格、条件模拟简介；第十章是地质统计学计算程序。

在编写过程中，编者力求做到基本概念清楚，数学推导严密，通俗易懂，逐步深入。为了使读者更好地掌握地质统计学的应用方法，书中较多地运用了我国地质统计学方面有关的科研成果作为实例。

本书在编写过程中得到张大顺副教授、许友志副教授的热情关怀和指导，并受到许多专家们的大力支持，特别是张幼蒂教授阅读了全部初稿，并提了许多宝贵意见，在此一并表示感谢。

由于时间紧，编者水平有限，书中难免有错误和不妥之处，恳请读者批评指正。

编　者

1990.2.

试读结束：需要全本请在线购买： www.ertongbook.com

目　　录

第一章 绪 论

地质统计学是近20多年创立并发展起来的一门新兴边缘学科。它开始主要是为解决矿床从普查勘探、矿山设计到矿山开采整个过程中各种储量计算和误差估计问题而发展起来的。地质统计学是由法国著名学者G.马特隆教授于1962年创立的。在初期主要由他所领导的法国地质统计学学派加以推广、应用并发展起来的。经过20多年的发展，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，扩大了其应用领域，并已由法国、南非及一些法语国家推广到几乎全世界。我国从1977年开始引进地质统计学，受到了广泛的重视。几年来，地质、冶金、煤炭等部门有不少单位陆续开展地质统计学的研究，并取得了初步进展。

从地质应用的角度看，地质统计学是数学地质的一个独立分支；从概率统计的角度看，它突破了经典统计学不考虑样品点空间分布的不足，因而地质统计学从某种意义上也可以说是应用统计学的一个新分支。

那么什么是地质统计学呢？地质统计学与经典统计学(数理统计)又有什么区别呢？下面我们看一个简单的例子：

若在某一勘探线上的5个钻孔中，测得其煤厚数据为 $x=(2, 4, 3, 1, 5)$（见图1-1(a)）；在另一勘探线上的5个钻孔中，测得煤厚数据为 $x=(1, 2, 3, 4, 5)$(见图1-1(b))。由数理统计分析这两组数据的均值和方差都是相同的。即：

$$\bar{x}=\frac{1}{5}\sum_{i=1}^{5}x_i=3$$

$$s=\sqrt{\frac{1}{5}\sum_{i=1}^{5}(x_i-\bar{x})^2}=1.4$$

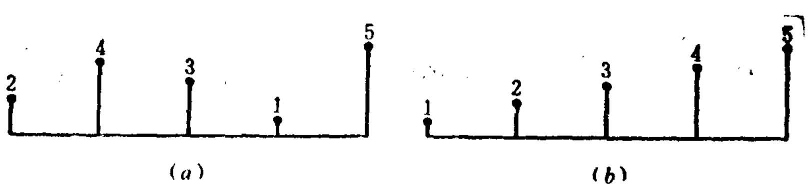


图1-1 钻孔煤厚数据示意图

但是，这两组煤厚数据所反映的煤层厚度的变异性（变化程度）是不一样的(见图1-1)。因此经典统计学不能区分这两组数据的煤厚变化性。这是因为经典统计学只考虑数值大小，不考虑数据的空间位置关系。用地质统计学方法来研究这两组数据，借助于变差函数，就能把这两组数据的变异性区别开来。这是由于地质统计学不仅考虑了地质变量的数值大小，而且还考虑地质变量的空间位置关系，使得插值计算或储量计算更加合理。

在此，我们初步尝到了地质统计学的甜头，至于它的理论体系、计算方法，我们将在后续的章节中陆续介绍。

为了更好地理解地质统计学，首先介绍一下地质统计学的产生背景。

第一节　地质统计学的产生

一、传统储量计算方法的评述

由于地质统计学是从矿石的品位估计和储量计算开始的，因此我们首先回顾一下储量计算的传统方法。

1．传统矿石储量计算公式为：

$$Q=V\cdot d \tag{1-1}$$

其中　Q——矿块矿石的储量；

V——矿块的体积；

d——矿石的体重。

而矿块的金属含量则是矿石储量与平均品位的乘积：

$$P=Q\cdot\bar{c}=V\cdot d\cdot\bar{c} \tag{1-2}$$

其中　P——矿块的金属含量；

$\bar{c}$——矿块的平均品位。

对于公式（1-2）中的矿石平均品位 $\bar{c}$，通常有两种方法求得：一种是当矿块各部位品位相近时，可用算术平均法由各剖面平均品位求矿块的平均品位；另一种方法是按样品在空间中的位置来加权计算，即所谓加权平均法。通常用的是距离 m 次反比法（图 1-2）。块段 B 的平均品位 $\bar{c}_B$ 是：

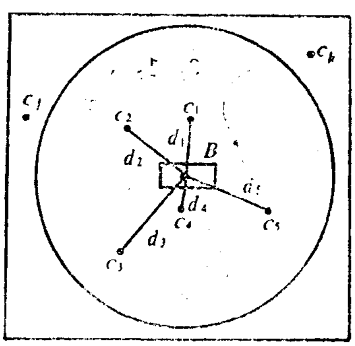


图1-2　距离（m次方）反比法示意图

$$\bar{c}_B=\sum_{i=1}^{n}\frac{c_i}{d_i^m}\Bigg/\sum_{i=1}^{n}\frac{1}{d_i^m} \tag{1-3}$$

其中　c_i——第 i 个信息样品点的品位值；

d_i——第 i 个信息样品点到待估块段 B 之间的距离；

m——距离的方次。

当 $m=1$ 时，式（1-3）称为距离反比法。

当 $m=2$ 时，式（1-3）称为距离平方反比法。

2．传统的煤炭储量的计算通式为：

$$Q=S\cdot M\cdot d\cdot\cos\alpha \tag{1-4}$$

其中　Q——煤炭储量；

S——面积；

M——厚度；

d——容重（或称体重）；

α——煤层倾角。

由此可见，储量是由矿体（煤层）的体积和容重相乘而得到的。其中容重 d 可用各种方法测得，由于它变化不大，故不作为储量计算研究的重点。在煤炭储量计算中，关键是准确地估计出煤层的厚度，进而求出煤层的体积，常用的煤炭储量计算方法主要有：

试读结束：需要全本请在线购买： www.ertongbook.com

算术平均法。将全区总面积乘以各钻孔煤厚的平均值来计算储量，其实质就是把整个复杂形状的煤层当作一个理想的，具有相同厚度的板状体（图1-3）。

地质块段法。实质上是把复杂形态的煤层当作一组密集的、大小不等的柱体（图1-4），计算各个柱体储量之和。

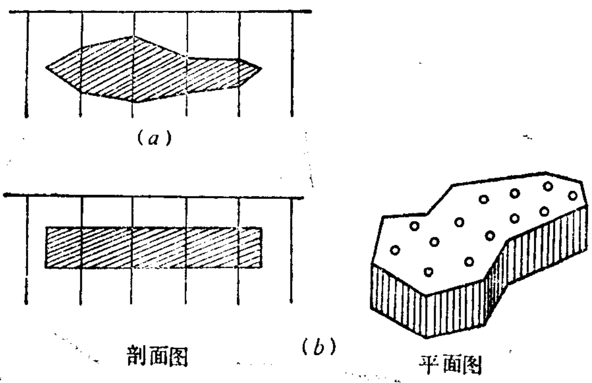


图1-3　用算术平均法计算储量示意图
a—地质剖面图上薄厚不均的煤层；
b—在储量计算时当成等厚的板状体

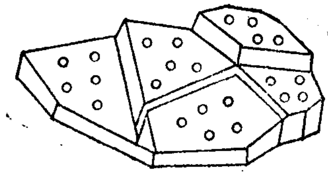


图1-4　地质块段法储量计算示意图

多角形法。在储量计算平面图上，以各个钻孔见煤点为中心，把煤层划分为若干个多角形块段，然后分别计算各块段储量。此法也称为最近地区法（图1-5）。

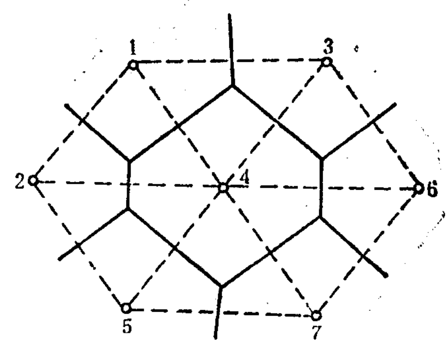


图1-5　用多角形法计算储量示意图

在回顾了传统储量计算方法以后，可以明显地看到传统储量计算方法存在不少问题：

1）简单地把某个钻孔的观测数据当作一个块段的平均数据值(例如“多角形法”)；或者是把周围样品数据的线性组合作为一个块段的值（例如“距离m次方反比法”）。这样对于复杂的矿体(或煤层)来说，一个样品的观测值不可能正好是影响范围（块段）的平均值。由于钻孔岩芯样品的承载（support，或叫“支撑”、“支集”，一般指样品或块段体积大小而言）小而块段的承载大，这二者的平均值是不能简单地加以等同看待的。表1-1引用了D.G.克立格在计算南非某金矿时，用传统多边形法估计品位与开采后的实际品位的对比情况表，从表中可以明显地看出上述算法的系统偏差。

表1-1　某金矿用传统方法估计品位和开采品位的对比

块段矿石等级 in-dwt①	矿　量 t	品位估值 in-dwt	开采后品位值 in-dwt	误　差
低级 91～149	110 000	124	150	－17%
中级 150～266	400 000	201	199	＋1%
高级 267～348	160 000	296	236	＋25%

（据D.G.Krige）

①注：金的衡量单位为 in-dwt＝0.165cm·g，表示垂直于矿层底面积为1平方英寸的柱内金的含量（in-dwt 是inch-pennyweight（吋-英钱）的缩写）。

2）未考虑矿体品位（煤层厚度）的空间变异性。如图1-6所示，假设方向V表示煤层的倾向方向，方向U表示煤层的走向方向。煤层厚度在U方向的变异性小于在V方向上的变异性，这时尽管样品Z_1, Z_2, Z_3，Z_4与待估块段S的距离相等，但在估计S的平均煤厚时，Z_2，Z_4的权就应比Z_1和Z_3的权大些。而传统的方法，对Z_1，Z_2，Z_3，Z_4只能赋予相同的权。

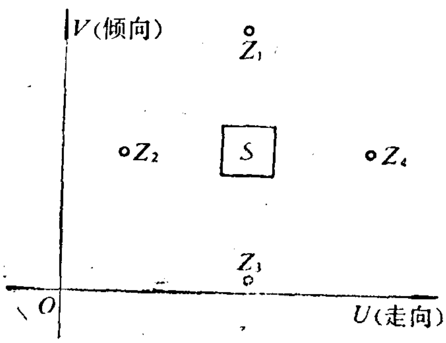


图1-6　煤厚的空间变异性对块段估值的影响

3）传统储量计算方法只能近似地给出一个矿体的平均品位(或平均厚度)，而不能给出品位（或厚度）在空间的变化性（即离散程度）的大小。如图1-1(*a*)，(*b*)所示，尽管其平均值及方差均相同，但其煤厚在两条勘探线上的变化性却不大一样。然而，对于这种空间变化程度的研究，对矿井设计是非常重要的。

4）没有衡量计算精度的方法和标准。传统检验精度的方法，是选用一个矿块，对它的储量用同一种方法计算两次或多次，视其相对误差的大小来估计；或者用两种方法计算同一块段的储量，用两者之间的相对误差来估计精度。故没有一种客观定量的衡量计算精度的方法和标准。

二、经典统计学的局限性

为了反映地质变量的空间变化性，一些地质学家和采矿工程师曾经使用了某些经典的概率统计方法来研究地质变量。由于地质变量并不总是纯粹的随机变量，因此直接用简单的统计方法解决复杂的地质问题，有一定的局限性。

在西方，最先具体应用统计学方法的是南非学者 H.S. 西舍尔（H.S.Sichel)。在H.S.西舍尔所用的估计量中存在着三个主要缺陷：

(1）作为背景的概率分布必须是对数正态分布；

(2）每次抽样必须是独立的，或说样本独立；

(3）不考虑样本点空间位置的差异，所有样本点都是同等重要的。

实际上，这些缺陷是带有一般性的。经典统计学用于研究地质变量，至少有以下几方面的不足之处：

1．经典统计方法在研究地质变量时，不考虑样品的空间分布。在统计钻孔岩芯煤厚时，煤厚的空间分布，决定着煤层的稳定程度。如图1-1所示，在(*a*)、(*b*)两勘探线中，各有五个钻孔，煤厚数据均为1m，2m，3m，4m，5m。但由于样品的空间分布位置不同，尽管它们的均值、方差都是一样的，煤层的稳定性也不相同。这个例子直观地说明了经典统计方法不能反映地质变量的空间变化性。

2．经典概率统计学的研究对象必须是纯随机变量，且都要服从一定的已知概率分布。而许多地质变量并不是纯随机变量，而是既有随机性又有结构性（指空间分布上某种程度的相关性或连续性）的变量。

3．经典统计学所研究的变量原则上要求可以无限次地重复试验或大量的观测，但地质变量则不行。例如在某个勘探区，从一个钻孔取得某层煤的样品，一旦取出后，一般就不可能在同一个地方取得该层煤的样品了。

4．经典统计学一般要求每次抽样必须是独立进行的，即要求样本(x_1，x_2，…，x_n)中各个 x_i($i=1$，2，…，n)都是相互独立的，但地质变量（如煤厚或矿石品位）就不一定独立，往往有某种程度的空间相关性。

为了寻求一种既能保持概率统计方法的有效性，又考虑到地质变量特点的新途径，便导致了地质统计学的产生。

三、地质统计学的诞生

南非地质学家、采矿工程师D.G.克立格，1951年根据南非金矿的具体情况，提出了计算矿产储量的方法，按照样品与待估块段的相对空间位置和相关程度来计算块段品位及储量，并使估计误差为最小。后来法国学者G．马特隆对克立格提出的方法进行了研究，认为D.G．克立格提出的方法是在考虑了矿化空间分布特征的基础上，合理地改进了统计学，是一种传统储量计算方法和统计学方法高度地结合起来的新方法。

为了解决在具有二重性（随机性和结构性）的地质变量的条件下使用统计方法的问题，1962年，G.马特隆在经过深入讨论和实践之后，提出了区域化变量(Regionalized Variable)的概念，产生了地质统计学(Geostatistics)。根据地质统计学理论，矿化特征可以用区域化变量的空间分布来表征。而研究区域化变量空间分布的主要数学工具是变差函数(Variogram；有的书上叫"半变异函数"，Semivariogram；二者只是在定义上相差常数系数$\frac{1}{2}$，并无本质区别)。

G．马特隆指出："从历史的角度来看，地质统计学和采矿业是同时出现的。就反映矿化空间特征而论，那些传统的方法仍保持着它们的全部价值，现代的理论成就决不是要否定这些方法，而是把它们当作出发点，并将其提到更高的科学水平。"地质统计学的诞生为在地质勘探和采矿工程中应用新的统计学方法开拓了前景。

第二节　地质统计学的现状及优点

一、地质统计学的现状

经过20多年的发展，目前地质统计学已初步形成了一套完整的理论体系，提出了一些重要的方法和技巧，编制了一套有实际应用价值的程序包，并迅速传播到世界各地。不过它仍然是一门年轻的学科，还正在处于蓬勃发展阶段。下面介绍有关方面的情况。

1．初步形成了一套完整的理论体系

1）线性平稳地质统计学。这是地质统计学的基础部分，包括基本概念[区域化变量]，基本工具[变差函数]，基本假设[二阶平稳假设和本征(intrinsic或译为"内蕴"）假设]，基本公式[估计方差，离散(dispersion或译为"离差"）方差，正则化公式]和基本方法[普通克立格法]等内容；

2）线性非平稳地质统计学：包括泛克立格法和K阶本征函数法等；

3）条件模拟：包括对矿床的条件模拟和对采矿过程的条件模拟等；

4）平稳非线性地质统计学，包括析取克立格法等；

5）储量参数的确定。

2．提出了一些重要的方法和技巧：如正态变形法、旋转薄片法等。

3．编制了一整套实际有效的程序包和程序库：包括变差图拟合程序包、地质统计学程序库、正态变形程序库等。

4．扩大了地质统计学的应用领域。

1）在地质、采矿方面：已从传统的估计各种金属、非金属矿床储量，发展到估计煤田和油田（原地）储量；还发展到应用于水文地质学、海洋地形测深学，还能解决区域地质以及地质制图等问题；制定合理的勘探方案（如合理的勘探网密度的确定、合理勘探程度、合理勘探精度及最优井位的确定等）。

2）其它应用领域：由于区域化变量是广泛存在于自然现象之中的，因而地质统计学还可应用于地质和矿业以外的领域，如森林资源的估计、农作物估产、大气降雨量估计以及环境保护学科中的大气污染估计等。总之，凡是属于对空间分布的数据进行最优线性无偏内插估计问题，都可以应用地质统计学。由此可见，地质统计学的应用领域是相当广泛的。

地质统计学的研究与应用已由法语国家扩展到世界各地，目前美国、加拿大、意大利、英国、澳大利亚、苏联、西班牙以及许多拉丁美洲国家都相继开展了地质统计学的研究与应用。法国矿业界已把地质统计学作为储量计算的标准方法。我国在1974年开始见到国外这方面的论文。1977年我国冶金、地质系统的不少单位开展了这方面的工作，并取得了初步进展。在煤炭系统沈阳煤矿设计院等单位应用地质统计学较早。目前全国已有不少的高等院校在本科生和研究生中开设地质统计学课程。

现在地质统计学已经成为国际性的学科，国际上已经召开了三次地质统计学大会，它在数学地质中占有的地位日益提高。

二、地质统计学的优点

1．什么是地质统计学

1962年G．马特隆曾给地质统计学下了较广泛的定义：“地质统计学就是随机函数的形式体系在勘探与估计自然现象上的应用。”随着地质统计学的发展，也有学者提出：“地质统计学是区域化变量理论在那些展布于空间、并显示出一定结构性和随机性的自然现象上的应用。”如果更具体地从各个侧面来描述，我们则可以说：“地质统计学是以变差函数作为基本工具，在研究区域化变量的空间分布结构特征规律性的基础上，综合考虑空间变量的随机性和结构性的一种数学地质方法。”

2．地质统计学的优点

地质统计学之所以能得到迅速发展和广泛应用，是因为它与传统的计算储量方法和经典概率统计方法在地质采矿的应用相比有一系列明显的优点。主要是：

1）地质统计学并不是将经典统计方法简单地用于地质、采矿领域，而是从地质、采矿的实际出发，对原有的数学理论与方法加以选择、创新，使之更有效地解决地质、采矿中的问题。

2）它最大限度地利用勘探工程所提供的各种信息。在使用克立格法估计矿床中某一块段变量的平均值时，充分考虑了待估块段与信息样品之间的空间位置关系、信息样品彼此之间的空间位置关系以及区域化变量空间分布的结构特征等方面的因素。

3）它不仅可以进行储量的整体估计，还可以进行储量的局部估计。在基本开采方法已确定的条件下，可以分别算出最小开采块段的品位和储量，使勘探报告提供的储量能

更好地满足矿山设计要求，有利于矿床勘探、矿山设计和矿山开采三个阶段的互相衔接。

4）估计精度高，至少可以避免系统误差。例如，美国亚利桑那州有一个矿山，应用传统方法的一个模型预测了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%。最后开采实际比例为59.4%和40.6%。可见地质统计学的估计精度之高。

5）能具体地给出估计精度的概念。地质统计学中的克立格方差，是度量估计精度的一个很好的尺度，而传统方法则无法给出估计精度。在这方面地质统计学显示了明显的优越性。实际上，地质统计学中的克立格法给出的估计，还是一种线性无偏最佳（最小方差）的内插估计。

6）应用地质统计学方法与应用电子计算机相结合，使得储量计算实现了计算机化，这样既可节省人力和时间，又可提高储量计算的质量。

7）它可用于计算机自动成图，还可分块拼图。由于克立格法是一种线性无偏最佳的内插估计，因而可以为计算机自动成图提供很有力的工具和方法。国外这方面已取得明显效果。在地质领域中，由于很多成果都要用各种图件表示，故用地质统计学方法进行计算机自动成图的研究还是大有可为的。

8）地质统计学中的条件模拟可以很好地再现地质变量的变化性（或曲线的波动性）。它不仅可以模拟矿体、矿床，还可以模拟矿山开采过程，这就为矿山设计、开采最优化提供了定量依据和基础。

由于地质统计学是从研究矿产储量计算和误差估计问题产生、发展起来的，自然它在这些方面的研究比较成熟。但是它在地质勘探中应用的前景也是很大的。我们可以利用它来建立矿床地质特征、勘探方法和储量精度三者之间的关系，从而研究出矿床的合理勘探方案。

第二章　线性代数及概率统计基本知识

第一节　线性代数的基本知识

一、矩阵及其基本运算

1．矩阵基本概念

在地质统计学中，矩阵是一种不可缺少的工具。例如，对于n个样品、m个变量的地质数据进行分析，并将每一样品的分析数据逐次记录如下：

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{matrix}$$

其结果就是一个包含了$(n\times m)$个数据的矩阵。

矩阵　$n\times m$个数a_{ij} $(i=1,\cdots,n,\ j=1,\cdots,m)$所排列成的矩形数表称为一个$n$行$m$列的矩阵，或简称为$n\times m$矩阵，并用$\boldsymbol{A}$表示

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

a_{ij}称为矩阵$\boldsymbol{A}$的第i行第j列的“元素”。$\boldsymbol{A}$有时记作$\boldsymbol{A}=[a_{ij}]_{n\times m}$。$i$称为行下标，$j$称为列下标。如果$m=n$时，则称矩阵$\boldsymbol{A}$为$n$阶“方阵”（或$n$阶矩阵）。

矩阵中，行下标与列下标相等的元素$(a_{11},a_{22},\cdots,a_{ii},\cdots,a_{nn})$的连线叫该矩阵的“主对角线”。除主对角线以外的元素都是0的矩阵称为“对角矩阵”。

2．矩阵运算

对于矩阵，可以施行下列基本代数运算：矩阵的加法、矩阵与数的乘法以及矩阵与矩阵的乘法。

1）矩阵的加法　两个$(n\times m)$矩阵相加就是将其全部元素对应相加。例如$\boldsymbol{A}=[a_{ij}]_{n\times m}$，$\boldsymbol{B}=[b_{ij}]_{n\times m}$则

$$\boldsymbol{A}+\boldsymbol{B}=[a_{ij}+b_{ij}]_{n\times m} \tag{2-1}$$

2）矩阵与数的乘法　数α与矩阵$\boldsymbol{A}$的积$\alpha\boldsymbol{A}$就是用α遍乘$\boldsymbol{A}$的全部元素所得到的$(n\times m)$矩阵，即

$$\alpha\boldsymbol{A}=[\alpha a_{ij}]_{n\times m} \tag{2-2}$$

例1　已知：

$$\boldsymbol{A}=\begin{bmatrix} 1 & 2 & 3 \\ -1 & 5 & 0 \end{bmatrix},\qquad \boldsymbol{B}=\begin{bmatrix} 1 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

则

试读结束：需要全本请在线购买：　www.ertongbook.com