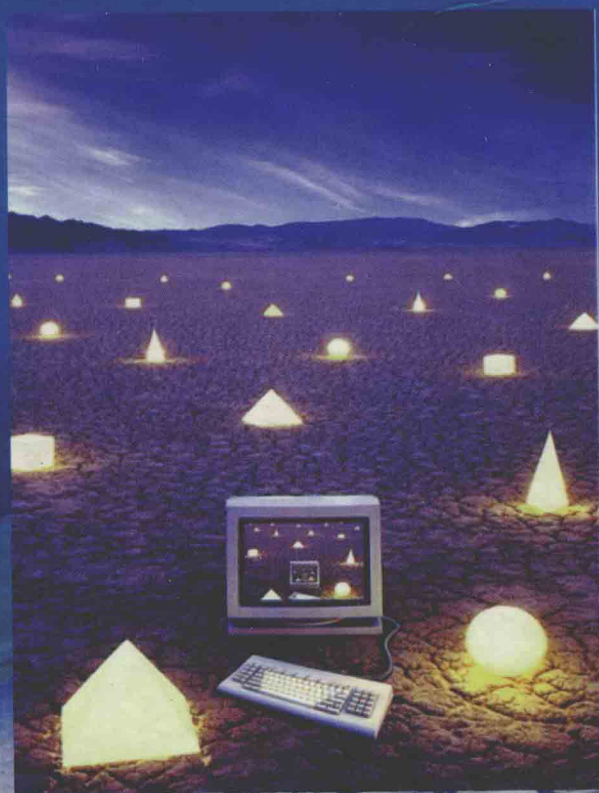


石广仁 著

地学中的 计算机应用新技术



石油工业出版社

PETROLEUM INDUSTRY PRESS

地学中的计算机应用新技术

石广仁 著

石 油 工 业 出 版 社

内 容 提 要

本书较全面地介绍地球科学领域中计算机应用的各种新技术,包括四大类:概率、统计、优化类;模糊、灰色类;神经网络、分形几何类和地质统计类。对于每一大类中的各种技术,分别给出其应用范围及条件、基本原理、计算方法和应用实例。在每个应用实例中,均给出它们的计算流程图、源程序、计算结果及其分析。书中提供的17个源程序达到通用化的水平,读者只要修改输入参数内容,就可将源程序变为自己的应用软件。

本书内容技术先进、实例典型、实用性强,具有广泛的使用和参考价值。可作为高等院校地学专业本科生和研究生的教材,也可供地学科技人员使用。其实,上述四大类新技术中,除地质统计类专门应用于地学外,其余三大类也可以应用于其它学科。

图书在版编目(CIP)数据

地学中的计算机应用新技术/石广仁著.

北京:石油工业出版社,1999.9

ISBN 7-5021-2745-3

I.地…

II.石…

III.计算机应用-地球科学-研究-新技术

IV.P-39

中国版本图书馆CIP数据核字(1999)第39494号

石油工业出版社出版

(100011 北京安定门外安华里二区一号楼)

石油工业出版社印刷厂排版印刷

新华书店北京发行所发行

787×1092毫米 16开本 19印张 464千字 印1—1500

1999年9月北京第1版 1999年9月北京第1次印刷

ISBN 7-5021-2745-3/TP·43

定价: 98.00元

前 言

俗话说：“上天容易，入地难”，是指人类或测量仪器所能抵达目的地的难度而言。这迫使地球科学工作者必须在使用极少量直接方法（如钻探）、大量间接方法（如地面仪器）的条件下，尽可能多地应用计算机技术，根据有限而十分宝贵的测量资料，由点到线、由线到面、由面到体地推测所研究的地下目标的整体状况，为发现与开发地下资源发挥最大限度的作用。

由于地下资源分布的非均质性、随机性及复杂性，人们采用的各种计算机技术经过实践发现：有的成功，有的失败。成功的技术限定于一定的应用范围及条件，失败的技术归因于客观条件所限。在近 40 年的探索应用中，地学中的计算机应用可归为三类：海量信息处理（如遥感、地震、测井等）、数值模拟（如盆地模拟、油藏模拟等）和若干计算机科学的技术。本书不论及前两类，只讨论第三类。在第三类中选择那些先进而实用的技术介绍给读者，内容分为如下四类：概率、统计、优化类；模糊、灰色类；神经网络、分形几何类和地质统计类。

在概率、统计、优化类（第一章）中，概率密度函数（Probability Density Function）、蒙特卡罗法（Monte Carlo Method）属概率论在地学中的应用；线性关系式（Linear Relation）和多项式（Polynomial）的最小二乘法（Least Square Method）、逐步回归法（Successive Regression）属数理统计在地学中的应用；线性规划（Linear Programing）的单纯形法（Simplex Method）实现优化属运筹学在地学中的应用。

在模糊、灰色类（第二章）中，模糊综合评判法（Fuzzy Integrated Decision）可用于地质目标的综合评价，属模糊数学（Fuzzy Mathematics）在地学中的应用；灰色预测法（Grey Prediction）可用于时间域或空间域的预测，灰色综合评判法（Grey Integrated Decision）可用于地质目标的优选，它们都属于灰色系统（Grey System）在地学中的应用。

在神经网络、分形几何类（第三章）中，人工神经网络（Artificial Neural Network）可用于时间域或空间域的预测，分形几何（Fractal Geometry）可用于地质特性分析，它们都属于人工智能（Artificial Intelligence）在地学中的应用。

在地质统计类（第四章）中，地质统计的预处理（Data Pre-processing of Geostatistics）、实验变差函数（Experimental Variogram）、用球状模型（Sphere Model）实现实验变差函数的优化拟合（Optimal Fitting of Experimental Variogram）、克立格法的检验（Checking of Kriging）和克立格法的应用（Application of Kriging）等五个部分，虽是在不同的节中介绍，由不同的程序表达，但相互是有机联系的，合起来就是“点”克立格法在地学中的应用。地质统计是传统的概率统计学与地质学相结合的产物，在空间上由已知推未知的应用中其效果明显优于第一章介绍的概率、统计。

对于上述四大类中的各种技术，分别给出其应用范围及条件、基本原理、计算方法和应用实例。在每个应用实例中，均给出它们的计算流程图、源程序（17 个程序）、计算结果及其分析，从而确保本书论及的各种技术的完整性和可操作性。读者根据所指出的应用范围及条件，就可选择合适的技术，应用于相应的领域，将会收到良好的应用效果；根据简述的基

本原理，就可迅速掌握其技术实质，把握住应用方向；根据详述的计算方法，就可以了解计算过程的总体结构与技术细节、方法的优点与弱点，也为迅速看懂源程序打下基础；根据以英文详尽注释的源程序（17个程序，参见表1），就可以迅速地应用到自己的工作中，也可以轻而易举地把它集成到自己的软件中；根据那些“麻雀虽小、但五脏俱全”的实例及其计算结果、结果分析，就可以举一反三地应用到自己的实际工作中，迅速得到所需的结果。总之，本书适用于方法研究人员、软件研制人员及实际应用人员，使他们都能收到立竿见影的效果。

表 1 源程序一览表

序号	源程序名	方 法	应 用 举 例	正 文	源程序及 计算结果
1	ProDenFun	概率密度函数	高勘探程度地区油藏的 概率密度函数的计算	第一章第 1 节	附录 1-1
2	PorVolTrap	蒙特卡罗法	圈闭孔隙体积的计算	第一章第 1 节	附录 1-2
3	LinFitting	最小二乘法 (线性关系式)	孔隙度与声波时差的线性关系式；孔隙度与埋深的指数关系式	第一章第 2 节	附录 1-3
4	OrPoFitting	最小二乘法 (多项关系式)	孔隙度与声波时差的多项关系式	第一章第 2 节	附录 1-4
5	RegresAnaly	逐步回归分析	探明储量与若干地质量的线性关系式	第一章第 2 节	附录 1-5
6	Simplex	单纯形法优化	钻探规划的最佳经济效益	第一章第 3 节	附录 1-6
7	FuzIntDec	模糊综合评判法	钻探目标的模糊综合评价	第二章第 1 节	附录 2-1
8	GreyPredic	灰色预测法	原油产量的预测；储层的预测	第二章第 2 节	附录 2-2
9	GreyIntDec	灰色综合评判法	钻探目标的灰色综合评价	第二章第 2 节	附录 2-3
10	BPtime	人工神经网络（BP 算法）的时间域推算	原油产量的预测	第三章第 1 节	附录 3-1
11	BPspace	人工神经网络（BP 算法）的空间域推算	XOR 问题；压裂酸化效果的预测	第三章第 1 节	附录 3-2
12	FractGeom	分形几何	断层分形特征与油气运聚的关系	第三章第 2 节	附录 3-3
13	GeoStatis1	坐标系变换	平面（X，Y）上离散分布的某地质量的地质统计预处理	第四章第 1 节	附录 4-1
14	GeoStatis2	实验变差函数	平面（X，Y）上离散分布的某地质量的实验变差函数	第四章第 2 节	附录 4-2
15	GeoStatis3	实验变差函数的最优拟合（球状模型）	平面（X，Y）上离散分布的某地质量的实验变差函数的最优拟合	第四章第 3 节	附录 4-3
16	GeoStatis4	克立格法的交叉检验（扇形区最近点法）	平面（X，Y）上离散分布的某地质量的克立格法的交叉检验	第四章第 4 节	附录 4-4
17	GeoStatis5	克立格法的应用（平面最近点法）	平面（X，Y）上离散分布的某地质量的克立格法的应用	第四章第 5 节	附录 4-5

随带提一下，上述的四大类技术中，除了地质统计类专门用于地学外，其余三类（概率、统计、优化类，模糊、灰色类和神经网络、分形几何类）均可以应用于其它学科。希望本书对计算机在地学中应用的普及与提高起着推动工作。

笔 者

1999.6.9

目 录

第一章 概率、统计、优化类	(1)
第1节 概率	(1)
一、应用范围及条件	(1)
二、基本原理	(2)
三、用概率密度函数预测未发现的资源	(3)
1. 计算方法	(3)
2. 应用实例：高勘探程度地区油藏的的概率密度函数的计算	(4)
(1) 计算流程图	(4)
(2) 源程序	(5)
(3) 计算结果	(5)
(4) 结果分析	(5)
四、用蒙特卡罗法计算某个地质量	(5)
1. 计算方法	(5)
2. 应用实例：圈闭孔隙体积的计算	(6)
(1) 计算流程图	(7)
(2) 源程序	(8)
(3) 计算结果	(8)
(4) 结果分析	(8)
第2节 统计	(8)
一、应用范围及条件	(8)
二、基本原理	(9)
三、用最小二乘法求一个地质量表达所研究的另一个地质量的关系式	(10)
1. 计算方法	(10)
2. 应用实例1：孔隙度与声波时差的线性关系式、孔隙度与埋深的指数关系式	(12)
(1) 计算流程图	(13)
(2) 源程序	(13)
(3) 计算结果	(13)
(4) 结果分析	(13)
3. 应用实例2：孔隙度与声波时差的多项关系式	(13)
(1) 计算流程图	(14)
(2) 源程序	(14)
(3) 计算结果	(14)
(4) 结果分析	(14)
四、用逐步回归法求多个地质量表达所研究的另一个地质量的关系式	(15)
1. 计算方法	(15)
2. 应用实例：探明储量与若干地质量的线性关系式	(18)
(1) 计算流程图	(19)
(2) 源程序	(19)

(3) 数据文件	(19)
(4) 计算结果	(19)
(5) 结果分析	(19)
第3节 优化	(21)
一、应用范围及条件	(21)
二、基本原理	(21)
三、用线性规划实现勘探开发方案的最优化	(23)
1. 计算方法	(23)
2. 应用实例：钻探规划的最佳经济效益	(27)
(1) 计算流程图	(29)
(2) 源程序	(29)
(3) 计算结果	(29)
(4) 结果分析	(29)
第二章 模糊、灰色类	(31)
第1节 模糊	(31)
一、应用范围及条件	(31)
二、基本原理	(31)
三、用模糊综合评判法确定地质目标的优劣	(32)
1. 计算方法	(32)
2. 应用实例：钻探目标的模糊综合评价	(32)
(1) 计算流程图	(33)
(2) 源程序	(33)
(3) 计算结果	(33)
(4) 结果分析	(33)
第2节 灰色	(34)
一、应用范围及条件	(34)
二、基本原理	(34)
三、用灰色预测法进行时空推算	(35)
1. 计算方法	(35)
2. 应用实例1：用灰色预测法进行原油产量的预测	(36)
(1) 计算流程图	(37)
(2) 源程序	(38)
(3) 计算结果	(38)
(4) 结果分析	(38)
3. 应用实例2：用灰色预测法进行储集层的预测	(39)
(1) 计算流程图	(39)
(2) 源程序	(40)
(3) 计算结果	(40)
(4) 结果分析	(40)
四、用灰色综合评判法确定地质目标的优劣	(40)
1. 计算方法	(40)
2. 应用实例：钻探目标的灰色综合评价	(43)
(1) 计算流程图	(45)

(2) 源程序	(46)
(3) 计算结果	(46)
(4) 结果分析	(46)
第三章 神经网络、分形几何类	(47)
第 1 节 神经网络	(47)
一、应用范围及条件	(47)
二、基本原理	(47)
三、BP 算法	(51)
1. 计算方法	(51)
2. 计算流程图	(55)
3. 一个简单实例: XOR 问题	(55)
四、用人工神经网络法进行时空推算	(57)
1. 应用实例 1: 用人工神经网络法进行原油产量的预测	(57)
(1) 计算流程图	(58)
(2) 源程序	(58)
(3) 计算结果	(58)
(4) 结果分析	(58)
2. 应用实例 2: 用人工神经网络法进行压裂酸化效果的预测	(59)
(1) 计算流程图	(60)
(2) 源程序	(60)
(3) 计算结果	(61)
(4) 结果分析	(61)
第 2 节 分形几何	(61)
一、应用范围及条件	(61)
二、基本原理	(61)
三、用分形几何法定量描述地质构造特征	(62)
1. 计算方法	(62)
2. 应用实例: 断层分形特征与油气运聚的关系	(63)
(1) 计算流程图	(64)
(2) 源程序	(65)
(3) 计算结果	(65)
(4) 结果分析	(65)
第四章 地质统计类	(66)
第 1 节 数据预处理	(67)
应用实例: 平面 (X, Y) 上离散分布的某地质量的地质统计预处理	(68)
(1) 计算流程图	(69)
(2) 源程序	(70)
(3) 数据文件	(70)
(4) 计算结果	(70)
(5) 结果分析	(71)
第 2 节 实验变差函数	(71)
一、应用范围及条件	(71)
二、基本原理	(71)

三、用实验变差函数描述地质量的分布特性	(72)
1. 计算方法	(72)
2. 应用实例：平面 (X, Y) 上离散分布的某地质量的实验变差函数	(76)
(1) 计算流程图	(76)
(2) 源程序	(76)
(3) 数据文件	(76)
(4) 计算结果	(77)
(5) 结果分析	(77)
第 3 节 实验变差函数的最优拟合	(78)
一、应用范围及条件	(78)
二、基本原理	(78)
三、用球状模型实现实验变差函数的最优拟合	(79)
1. 计算方法	(79)
2. 应用实例：平面 (X, Y) 上离散分布的某地质量的实验变差函数的最优拟合	(82)
(1) 计算流程图	(83)
(2) 源程序	(84)
(3) 数据文件	(84)
(4) 计算结果	(84)
(5) 结果分析	(84)
第 4 节 克立格法的交叉检验	(85)
一、应用范围及条件	(85)
二、基本原理	(85)
三、用扇形区最近点法实现克立格法的交叉检验	(87)
1. 计算方法	(87)
2. 应用实例：平面 (X, Y) 上离散分布的某地质量的克立格法的交叉检验	(90)
(1) 计算流程图	(90)
(2) 源程序	(91)
(3) 数据文件	(91)
(4) 计算结果	(92)
(5) 结果分析	(92)
第 5 节 克立格法的应用	(92)
一、应用范围及条件	(92)
二、基本原理	(92)
三、用平面最近点法实现克立格法的应用	(93)
1. 计算方法	(93)
2. 应用实例：平面 (X, Y) 上离散分布的某地质量的克立格法的应用	(95)
(1) 计算流程图	(96)
(2) 源程序	(97)
(3) 数据文件	(97)
(4) 计算结果	(97)
(5) 结果分析	(97)
参考文献	(98)
附 录 各源程序及实例计算结果	(99)

附录 1-1	高勘探程度地区油藏的概率密度函数的计算	(99)
附录 1-2	用蒙特卡罗法计算圈闭的孔隙体积	(103)
附录 1-3	用最小二乘法求孔隙度与声波时差的线性关系式、孔隙度与埋深的指数 关系式	(117)
附录 1-4	用最小二乘法求孔隙度与声波时差的多项关系式	(122)
附录 1-5	用逐步回归法求探明储量与若干地质量的线性关系式	(128)
附录 1-6	用线性规划实现钻探规划的最佳经济效益	(136)
附录 2-1	用模糊综合评判法进行钻探目标的综合评价	(165)
附录 2-2	用灰色预测法进行原油产量、储集层的预测	(169)
附录 2-3	用灰色综合评判法进行钻探目标的综合评价	(179)
附录 3-1	用人工神经网络法进行原油产量的预测	(187)
附录 3-2	用人工神经网络法进行压裂酸化效果的预测	(204)
附录 3-3	用分形几何法定量描述断层分形特征与油气运聚的关系	(222)
附录 4-1	平面 (X, Y) 上离散分布的某地质量的地质统计预处理	(227)
附录 4-2	平面 (X, Y) 上离散分布的某地质量的实验变差函数	(238)
附录 4-3	平面 (X, Y) 上离散分布的某地质量的实验变差函数的最优拟合	(252)
附录 4-4	平面 (X, Y) 上离散分布的某地质量的克立格法的交叉检验	(262)
附录 4-5	平面 (X, Y) 上离散分布的某地质量的克立格法的应用	(277)

第一章 概率、统计、优化类

本章介绍概率论、统计学和最优化在地球科学领域中应用的先进且实用的计算机技术。每项技术均给出其应用范围及条件、基本原理、计算方法和应用实例。在应用实例中，又给出计算流程图、源程序、计算结果以及结果分析。应用实例虽小，但反映了计算的全过程，便于读者迅速理解并掌握所应用的技术。计算流程图及源程序虽指某个应用问题，但实际上是通用的，只要在计算流程图中将原来应用问题的内容改为读者自己的应用问题，在源程序中按“INPUT DATA”所提示的注释来修改输入参数，在两个注释行“…… change from application …… top”与“…… change from application …… bot”之间按读者自己的应用问题进行修改，然后编译、运行，立即得到读者自己所需的计算结果。

在第1节（概率）中介绍了两个应用实例：①高勘探程度地区油藏的**概率密度函数**（Probability Density Function）的计算，以例说明如何使用概率密度函数预测未发现的资源；②圈闭孔隙体积的计算，以例说明如何使用**蒙特卡罗法**（Monte Carlo Method）计算某个地质量。

在第2节（统计）中介绍了三个应用实例：①孔隙度与声波时差的线性关系式，以例说明如何使用**最小二乘法**（Least Square Method）求一个地质量表达所研究的另一个地质量的**线性关系式**（Linear Relation）；②孔隙度与声波时差的多项关系式，以例说明如何使用**最小二乘法**（Least Square Method）求一个地质量表达所研究的另一个地质量的**多项关系式**（Polynomial）；③探明储量与若干地质量的线性关系式，以例说明如何使用**逐步回归法**（Successive Regression）求多个地质量表达所研究的另一个地质量的关系式。

在第3节（优化）中介绍了一个应用实例——钻探规划的最佳经济效益，以例说明如何使用**线性规划**（Linear Programing）的**单纯形法**（Simplex Method）实现勘探开发方案的最优化。

第1节 概 率

一、应用范围及条件

1. 应用范围

常用的方面有如下两个：

（1）用概率密度函数预测未发现的资源

假设某地区已发现规模不等的资源，据此制作该资源的概率密度函数。在这个函数曲线上，可以得到横坐标各区间的概率（概率密度函数值×区间长度），此值减去对应区间上已发现资源的概率，所得差值乘以总发现个数就是这个区间上的未发现资源的个数。

（2）用蒙特卡罗法计算某个地质量

假设一个地质量计算公式中有若干参数，每个参数具有个数不等的观察值，如何计算这个地质量？最简单最粗糙的方法是先计算各参数的平均观察值，然后代入计算公式，得到该地质量的平均计算值。然而，地下参数是随机分布的，地质量应以随机分布的概率形式表示

出来,才能客观地反映这个地质量。解决的办法是:根据观察值,求出各参数的概率分布函数;再用蒙特卡罗法求出该地质量的概率分布函数,即得到了各种概率下的地质量。一般输出 0.95, 0.50 及 0.05 三种概率所对应的地质量,并以概率为 0.50 所对应的地质量为所求。

2. 应用条件

各参数的观察数量足够多、观察值较精确、观察位置分布较均匀。

二、基本原理

1. 事件

观察地质现象,会发现如下三种情况:

①在一定条件下必然会发生的称为必然事件。例如,烃源岩的成熟度均等于或大于生油门限,干酪根必然生油。

②在一定条件下不可能发生的事件称为不可能事件。例如,烃源岩的成熟度均小于生油门限,干酪根不可能生油。

③在一定条件下,可能发生也可能不发生的事件称为随机事件。例如,烃源岩的成熟度有的等于或大于生油门限,有的小于生油门限,则随意抽查一个烃源岩时就出现干酪根可能生油、可能不生油的情况。假设某区有不少井,各井有若干个烃源层,当大量进行上述抽查时会呈现出生油概率的稳定性,即该区的生油概率(生油的烃源岩个数/抽查的烃源岩个数)趋近于一个稳定值。这说明生油这个随机事件发生的可能性的,是事件本身所固有的不以人们的主观意愿而改变的一种属性。事件的这种属性是人们对它发生的可能性的进行度量的客观基础。

概率论就是人们研究随机现象统计规律性的一门科学。

2. 概率

(1) 单事件概率

事件 A 出现的概率定义为

$$P(A) = \frac{A \text{ 出现的次数}}{\text{观察事件 A 是否出现的总数}} \quad (1-1)$$

(2) 多事件概率

设多个事件 $A_1, A_2, \dots, A_n$ 相互独立,则它们同时出现的概率定义为

$$\begin{aligned} P(A_1 A_2 \cdots A_n) &= P(A_1) \cdot P(A_2) \cdots P(A_n) \\ &= \prod_{i=1}^n P(A_i) \end{aligned} \quad (1-2)$$

3. 概率的密度函数和分布函数

(1) 概率密度函数 $p(x)$

$$\int_{-\infty}^{+\infty} p(x) dx = 1 \quad (1-3)$$

例如,均匀分布的概率密度函数为

$$p(x) = \begin{cases} \frac{1}{b-a} & \text{当 } a \leq x \leq b \\ 0 & \text{其它情况} \end{cases}$$

从图 1-1 可见,概率密度函数表示了概率在各个随机变量上的值;概率密度函数与随机变量轴(横坐标轴)所围的面积等于 1,即符合式(1-3)的定义。

(2) 概率分布函数 $F(x)$

$$F(x) = \int_{-\infty}^x p(x) dx \quad (1-4)$$

$$F'(x) = p(x) \quad (1-5)$$

例如，均匀分布的概率分布函数为

$$F(x) = \begin{cases} 0 & \text{当 } x \leq a \\ \frac{x-a}{b-a} & \text{当 } a \leq x \leq b \\ 1 & \text{当 } x \geq b \end{cases}$$

易见，均匀分布的概率密度函数和概率分布函数之间的关系是满足式 (1-4) 和式 (1-5) 这两个条件的。

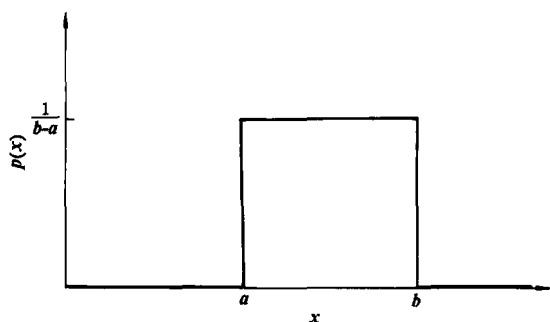


图 1-1 均匀分布的概率密度函数

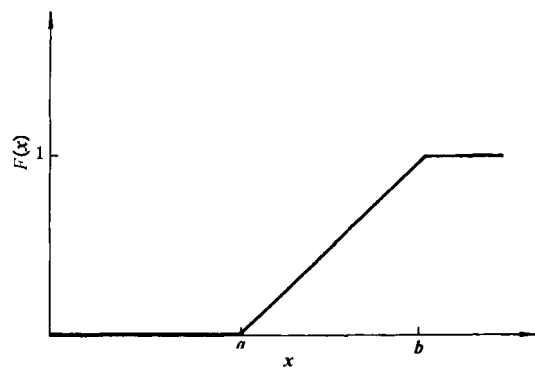


图 1-2 均匀分布的概率分布函数

从图 1-2 可见，概率分布函数表示了概率沿着随机变量轴（横坐标轴）的累积值，即式 (1-4) 的定义。

三、用概率密度函数预测未发现的资源

1. 计算方法

假设某高勘探程度地区已发现规模不等的油藏，它们的储量分别为 $q_1, q_2, \dots, q_n$ 。据此制作油藏的发现密度图：它们落在横坐标 $[q_i^*, q_{i+1}^*]$ 的油藏有 n_i 个 ($i=1, 2, \dots, m$)，则该区间上的纵坐标为 $n_i / [n (q_{i+1}^* - q_i^*)]$ ， $q_i^* < q_{i+1}^*$ ， $n = \sum_{i=1}^m n_i$ ，参见图 1-3。图 1-3 上有一条锯齿状的发现密度曲线，或称它为前面式 (1-3) 定义的概率密度函数的“毛坯”。根据“毛坯”的形态，选择一个合适的概率密度函数。

关于油藏预测中的概率密度函数，这里介绍如下的对数正态分布

$$p(q) = \frac{1}{\sqrt{2\pi}\sigma q} \exp\left\{-\frac{[\ln(q) - \mu]^2}{2\sigma^2}\right\} \quad (1-6)$$

式中

$$\mu = \frac{1}{n} \sum_{i=1}^n \ln(q_i) \quad (1-7)$$

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n [\ln(q_i) - \mu]^2 \quad (1-8)$$

确定了式 (1-6) 所示的概率密度函数，就将它画在图 1-3 上。

对于一个低、中勘探程度地区，已发现了一些油藏，如何测出尚未发现的油藏规模及其

个数呢？可用上述的对数正态分布（用类比法确定本地区的 μ 和 σ ），在图上画出该地区的储量的概率密度函数，再在该图上画出已发现油藏的发现密度图。可以肯定，这些发现密度的“条柱”不能充满概率密度函数曲线下方的面积。这些未充满部分便是尚未发现的油藏，由图可以确定它们的规模及其个数，参见图 1-4。

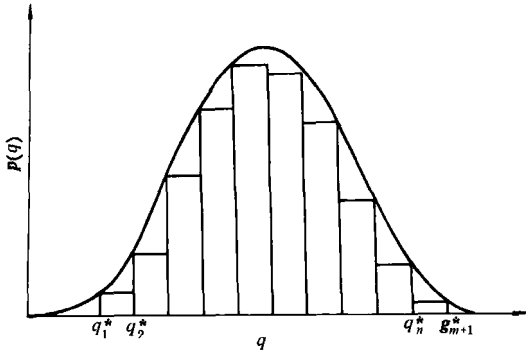


图 1-3 高勘探程度地区油藏的发现密度及概率密度函数

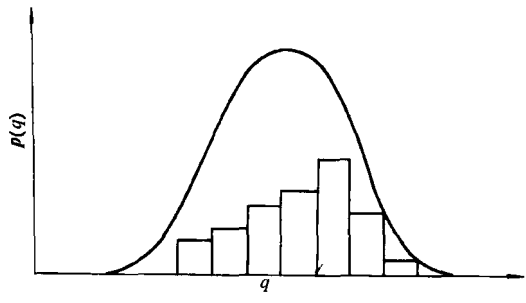


图 1-4 低、中勘探程度地区油藏的发现密度及概率密度函数

2. 应用实例：高勘探程度地区油藏的概率密度函数的计算

某盆地是高勘探程度地区，已发现 40 个油藏，它们的地质储量 q_k ($k=1, 2, \dots, 40$) 分别为：1.5, 2.3, 2.7, 3.2, 3.4, 3.6, 3.8, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 8.4, 8.5, 8.6, 9.5, 9.6, 10.5 (10^6t)。各个统计区间的端点值 q_i^* ($i=1, 2, \dots, 11$) 分别为：1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11。求该盆地油藏的概率密度函数。

(1) 计算流程图

参见图 1-5。

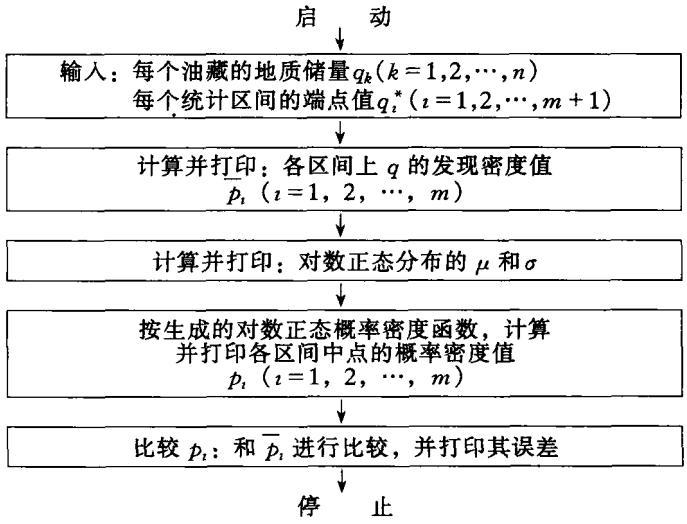


图 1-5 概率密度的计算流程图（程序名：ProDenFun）

(2)·源程序

参见附录 1-1 的 1-1-1。

(3) 计算结果

参见附录 1-1 的 1-1-2。

(4) 结果分析

采用上述的计算方法,算出对应于各统计区间的发现密度值:0.025, 0.05, 0.1, 0.15, 0.2, 0.175, 0.15, 0.075, 0.05, 0.025。可以验证,这 10 个数之和等于 1。这就是形如图 1-3 所示的锯齿状的发现密度曲线。采用式 (1-6) 表达的对数正态分布。按式 (1-7),算出 $\mu=1.713$;按式 (1-8),算出 $\sigma=0.5593$ 。将 μ 和 σ 代入式 (1-6),算出对应于各统计区间中点的概率密度函数值:0.03087, 0.1033, 0.1452, 0.1478, 0.1297, 0.1054, 0.08225, 0.06273, 0.04728, 0.03544。可以验证,这 10 个数之和等于 0.88997,不等于 1,这是由于总统计范围 $[1, 11]$ 的左、右端外没有进行统计的缘故,否则应等于 1,不过这对本问题的研究影响不大。这就是形如图 1-3 所示的概率密度函数。所以,该盆地油藏的概率密度函数为

$$p(q) = \frac{1}{1.4q} \exp \left\{ -\frac{[\ln(q) - 1.713]^2}{0.6256} \right\}$$

10 个发现密度值与 10 个概率密度值之间的均方差为 0.04424。

四、用蒙特卡罗法计算某个地质质量

1. 计算方法

假设一个地质质量计算公式中有若干参数,每个参数具有个数不等的观察值。下面逐步叙述这个地质质量的计算过程。

(1) 求各参数的概率分布函数

假设某参数的观察值为 $X_1, X_2, \dots, X_n$ 。据此制作该参数的累积频率分布图:落在横坐标 $[x_i^*, x_{i+1}^*]$ 的观察值有 n_i 个 ($i=1, 2, \dots, m$),则该区间上的纵坐标值为 n_i/n , $n = \sum_{i=1}^m n_i$,显然这是非累积的频率分布图;若将纵坐标值改为 $\sum_{k=1}^m n_k/n$ (易见 $i=1$ 时纵坐标值为 1),则得到累积频率分布,参见图 1-6。图 1-6 上有一条锯齿状的累积概率分布曲线,或称它为前面式 (1-4) 定义的概率分布函数的“毛坯”。

必须注意,若已知某参数的概率密度函数 $p(x)$,在计算时利用式 (1-5) 的积分形式 $F(x) = \int p(x)dx$ 将概率密度函数 $p(x)$ 值化为概率分布函数 $F(x)$ 值。后面的源程序就是这样设计的。

一个参数的概率分布函数确定了,再确定第 2 个参数的概率分布函数,直至各个参数都有各自的概率分布函数。

(2) 用蒙特卡罗法求地质质量

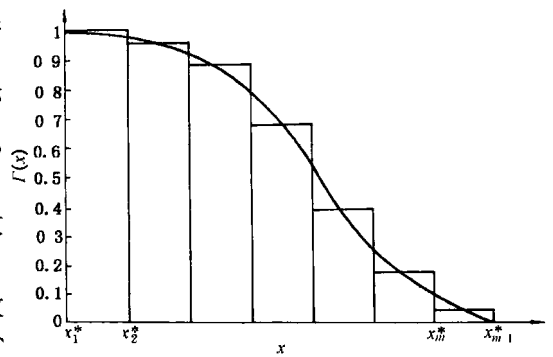


图 1-6 参数的累积频率分布及
概率分布函数

①伪随机数 γ_k 的生成。

采用混合同余法生成 $[0, 1]$ 区间上均匀分布的伪随机数, 递推公式如下^[1]

$$\left. \begin{aligned} \gamma_k &= X_k / M \\ X_{k+1} &= \alpha X_k + \beta \quad (\text{mod } M) \\ (k &= 1, 2, \dots) \end{aligned} \right\} \quad (1-9)$$

式中 X_k ——第 k 个伪随机数;

X_{k+1} ——第 $k+1$ 个伪随机数;

α ——乘子系数, 可取为 $5^5 = 3125$;

β ——增量, 可取为 3, 或 7, 或 11, 或 17;

M ——模, 可取为 $2^{19} = 524288$;

mod M ——按模式 M 的取余运算, 例如: $\alpha X_k + \beta = 1572874$, 除以 M (即 524288), 得到商为 3, 余为 10, 则 $X_{k+1} = 10$;

γ_k —— $[0, 1]$ 上的第 k 个伪随机数。

式 (1-9) 中, 初值 X_1 可取为 23, 或 11, 或 19, 或 37。 X_1 的 4 个值与上述 β 的 4 个值是对应的, 形成 4 组配对值, 应用中可选择任一组配对值。本书源程序中采用 $\beta = 3$, $X_1 = 23$ 。

②求对应于伪随机数 γ_k 的参数值 x_k 。

根据式 (1-9) 求得的 γ_k , 在参数的概率分布函数图 (参见图 1-6) 上求出对应的参数值 x_k 。由于 $F(x)$ 曲线是以离散点形式存放于计算机, 所以 r_k 落在纵坐标轴上某两个离散点之间, 所求的 x_k 落在横坐标轴上对应的两个离散点之间, 可用线性插值求出 x_k 。一个参数值确定了, 再确定第 2 个参数值, 直至各个参数值都确定。

③求对应于伪随机数 γ_k 的地质质量 G_k 。

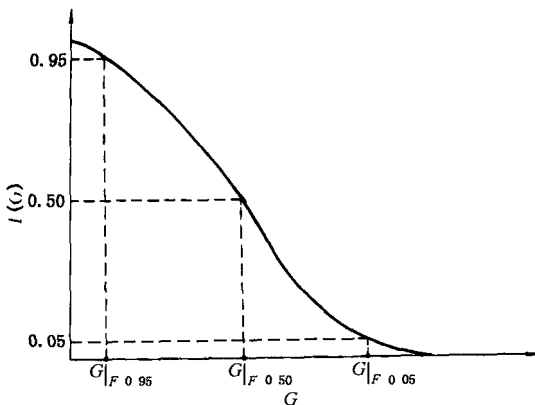


图 1-7 地质质量的概率分布函数

将求出的各参数值代入地质质量计算公式, 得到对应于伪随机数 γ_k 的地质质量 G_k 。

$k=1$ 时, 执行上述的三步计算, 得到 G_1 ;
 $k=2$ 时, 重复上述的三步计算, 得到 G_2 ;
……。生成的伪随机数个数, 也就是抽样的次数。抽样的次数在理论上越大越好, 但没有必要无限大, 只要由 γ_k 和 G_k ($k=1, 2, \dots$) 组成的地质质量概率分布函数的曲线形态达到稳定为止。一般地说, 当统计区间数为 100 时, 抽样次数可选为 500~5000 次。将每次抽样的 γ_k 和 G_k 画在图上, 最终得到该地质质量的概率分布函数, 参见图 1-7。

在图 1-7 上, $G|_{F=0.95}$ 和 $G|_{F=0.05}$ 分别表示在大累积概率和小累积概率下的地质质量值, $[G|_{F=0.95}, G|_{F=0.05}]$ 表示地质质量的可能范围。 $G|_{F=0.50}$ 处于 $G|_{F=0.95}$ 与 $G|_{F=0.05}$ 之间, 是最大可能的地质质量值, 也是所求的。

2. 应用实例: 圈闭孔隙体积的计算

在圈闭资源量计算中, 必须计算圈闭的孔隙体积, 公式为