文章编号:1007 - 3701(2005)03 - 0035 - 03

矿床储量规模分布的分形模型及其分维的数理分析

万 丽[1,2],邓 军[1],王庆飞[1],高邦飞[1]

(1. 中国地质大学地球科学与资源学院,北京 100083 2. 广州大学数学与信息科学学院,广州 510405)

摘要:矿床储量规模预测在矿产勘查和评价中占有十分重要的位置,而其预测的最主要手段是建立矿床储量规模分布的数学模型。根据矿床储量的统计自相似性,应用概率统计方法建立了矿床储量规模的分形模型。同时,通过模型及其分维值的数理分析,得出不同分维值的地学解释,讨论了分维的线性与非线性最小二乘回归估计法。

关 键 词:矿床储量;规模分布;分形模型;分维

中图分类号:P624.7 文献标识码:A

在矿产资源勘查和定量评价中,对某一成矿远景区的资源量和矿床数进行预测是前期勘查的重要工作,是政府对矿业开发宏观决策的重要依据[1]。目前,矿床储量规模预测最主要的手段是建立矿床储量规模分布的数学模型[2]。近年来,利用分形模型进行储量预测较为普遍[3,4],但对矿床储量规模的分形模型及分维的理论基础研究涉及甚少,对该模型的数理分析较缺乏。作者根据矿床储量的统计自相似性,应用概率统计方法建立了矿床储量规模的分形模型。同时,通过模型及其分维值的数理分析,得出不同分维值的地学解释,讨论了分维的线性与非线性最小二乘回归估计法,对这两种方法进行了比较。

1 模型建立

设某矿区分布有若干矿床 N,用矿床储量 r 作为度量尺度,则矿床储量为随机变量,设随机变量 R 的分布密度函数为 $f(r)$,储量大于 r 的概率记为 $P(r)=P(R\geqslant r)$,则

$$P(R\geqslant r)=\int_r^{+\infty} f(r)\,dr \qquad (1)$$

由于储量规模分布具有自相似性[5],即标度不变性,故满足

$$P(\lambda r)\propto P(r)G(\lambda) \qquad (2)$$

式中 λ 是变换常系数;$G(\lambda)$ 是与 r 无关的,只与 λ 有关的函数,G. Korvin (1992)(据文献[6])证明,满足(1)式的只可能是幂函数,考虑到 $P(r)\leqslant 1$,所以幂指数取负号即

$$P(r)\propto r^{-D},D>0$$

在实际问题中,通常用频率近似代替概率。

设已知区域的矿床总数为 N_0,大于储量 r 矿床累积数是 $N(\geqslant r)=N(r)$,则矿床累积数的频率是

$$P(r)=N(\geqslant r)/N_0 \qquad (3)$$

由此有 $N(\geqslant r)/N_0\propto r^{-D}$

从而得矿床储量规模的统计分形模型

$$N(r)=N(\geqslant r)=Cr^{-D} \qquad (4)$$

其中 $r>0$,$D>0$,C 为常数,D 称为分维数。

2 模型的数学性质

由(4)式,$N(r)$ 对 r 求导数为:

一阶导数 $N'(r)=-CDr^{-D-1}<0$,得 $N(r)$ 是 r 的递减函数

二阶导数 $N''(r)=CD(D+1)r^{-D-2}>0$,得

收稿日期:2005 - 01 - 25
基金项目:教育部科学技术研究重点项目(01037 和 03678);国家自然科学基金重点项目(40172036).
作者简介:万丽,女,副教授,博士生,现主要从事数学地质研究工作.

$N(r)$ 是 r 的凸函数。

由(1)得概率密度函数 $f(r) = -CDr^{-D-1}$

又由(3) $P(r) = P(R \geqslant r) = Cr^{-D}/N_0$

此时要求 $Cr^{-D} \leqslant N_0$,即 $r \geqslant (C/N_0)^{1/D}$

由此可得随机变量 R 的分布函数为

$F(r) = P(R \leqslant r) = 1 - (Cr^{-D}/N_0)$

数学期望为 $E(r) = CD/(D-1)$ $D>1$

方差 $Var(r) = \dfrac{CD[(D-1)^2 - CD(D-2)]}{(D-1)^2(D-2)}$, $D>2$

根据条件概率公式,有

$$P(R \leqslant r \mid R \geqslant r') = 1 - (r/r')^{-D} \qquad (5)$$

其中 $r \geqslant r' \geqslant (C/N_0)^{1/D} > 0$

(5)式说明,对任意常数 $k>0$,下式成立

$P(R \leqslant r \mid R \geqslant r') = P(R \leqslant kr \mid R \geqslant kr')$

即满足截尾条件下的分形不变性。

3 分维数的数理分析

又由(3)式可得 $r = (N(r)/C)^{-1/D}$

当 $N(r) = 1$,即储量大于 r 的矿床数为 1 时,$r = (1/C)^{-1/D}$ 应为最大储量,记为 r_1,即

$r_1 = (1/C)^{-1/D} = C^{1/D}$

当 $N(r) = N_0$ 时,$r = (N_0/C)^{-1/D}$ 应为最小储量,记为 r_n,即 $r_n = (N_0/C)^{-1/D} = (C/N_0)^{1/D}$

由此得 $r_1/r_n = N_0^{1/D}$ 和 $D = \ln N_0/(\ln r_1 - \ln r_n)$

当区域矿床数 N_0 一定时,D 越大,$\ln r_1 - \ln r_n$ 越小;D 越小,$\ln r_1 - \ln r_n$ 越大。特别:

①当 $D=1$ 时,有 $r_1 = C^{1/D} = C$,$r_1/r_n = N_0$,即 $r_1 = N_0 r_n$,说明最大储量恰好为最小储量与矿床总数之积,即满足 Zipf 定律的原始形式,矿床储量的各级值将分别为[4]:$C, C/2, C/3, L, C/n_0$。Zipf 定律在矿床储量预测中已有许多应用[7~9],特别是在前期勘查工作中,对基础地质研究程度较低的成矿远景区的储量预测。

②当 $D>1$ 时,$r_1/r_n = N_0^{1/D} < N_0$,说明矿床储量规模的分布相对比较集中,中间储量的矿床比较发育;D 越大,则 r_1/r_n 越小,储量等级就越集中。

③当 $0<D<1$ 时,$r_1/r_n = N_0^{1/D} > N_0$,说明矿床储量规模的分布差异较大,分布相对比较分散;D 越小,则 r_1/r_n 越大,储量等级差越大。

④当 $D \to \infty$ 时,即 $r_1/r_n \to 1$,说明矿床储量规模趋于一致,规模的分布是均匀的。

⑤当 $D \to 0$ 时,即 $r_1/r_n \to \infty$,说明区域内只有一个矿床,为绝对的首位型的分布。

后两种情况在现实中一般很难出现。

4 分维数的确定方法

4.1 一元线性回归法

首先,将观测数据($N(r_1), N(r_2), L, N(r_n)$)和($r_1, r_2, L, r_n$)绘在双对数坐标纸上,如果其散点大致分布在一条直线上,则直线的斜率便是分维数 D 的估计值。理论上,将(3)式两边取对数化为一元线性回归方程,即

$\lg N(r) = \lg C - D \lg r$

令 $y_i = \lg N(r_i)$, $x_i = \lg N(r_i)$ $i = 1, 2, L, n$

用最小二乘法得求出斜率 D 的估计值为:

$$\hat{D} = \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$,即分维值。

如果其散点大致分布在两条直线上,则需分段拟合,用上述方法分别确定每段直线的斜率即分维数 D_1、D_2 时,用最优化方法确定分界点,使得各区间拟合直线的剩余平方和 e_i ($i = 1, 2$) 的总和为最小。这种情况在现实中是常出现的,分界点常常是异常值的下限,是进行储量规模划分的重要参考依据。

4.2 非线性回归法

非线性回归法是将观测数据($N(r_1), N(r_2), L, N(r_n)$)和($r_1, r_2, L, r_n$)代入非线性回归方程(3),得观测方程

$N(r_i) = Cr_i^{-D} + \triangle_i$ $i = 1, 2, L, n$

其中 $\triangle_i$ 为观测误差。

设 $Q(C, D) = \sum_{i=1}^{n} [N(r_i) - Cr_i^{-D}]^2$

令 $\dfrac{\partial Q(C,D)}{\partial C} = 0$, $\dfrac{\partial Q(C,D)}{\partial D} = 0$

由于上述方程不易直接求出 C、D 的表达式,所以一般采用迭代方法求解[10]。通过迭代可得到非线性最小二乘估计参数 $\hat{C}$、$\hat{D}$。

需要说明的是,数学上迭代求解的方法有多种,常用的有 Newton 法、信赖域法、拟 Newton 法、最速下降法、Gauss – Newton 等,需要根据具体模型的非线性强度和约束条件,以及精度要求来确定[10]。

5　结语

分维数 D 是矿床储量规模分布的定量判据,利用分维数 D 的大小,可进行靶区优选以及有利控矿条件的确定;同时,对同一地区不同时期的分维值的比较,可动态分析矿床储量分布的变化规律。

分维值的确定有线性回归和非线性回归两种,由于一元线性回归法估计参数较简单,在实际工作中应用较多,但结果可能不正确,这是因为,对非线性强度较强的非线性模型,线性近似将产生大于观测误差的模型误差,所以,从理论上,应该采用非线性回归法,用非线性回归法估计出的参数,在拟合分形模型时更精确,即剩余平方和更小,而且不必受取对数条件要求的限制。

参考文献:

[1] 赵鹏大. 矿产勘查理论与方法[M]. 武汉:中国地质大学出版社,2001.
[2] Donald,Singer. 资源定量评价发展方向展望[J]. 地球科学,2001,26(3):152—156.
[3] 魏民,赵鹏大,刘红光,等. 中国岩金矿床品位 – 吨位模型研究[J]. 地球科学,2001,26(2):176—179.
[4] 刘庆生,燕守勋,赵善仁. 齐波夫金矿资源量预测[J]. 地质与勘探,1999,35(4):33—35.
[5] 申维. 分形混沌与矿产预测[M]. 北京:地质出版社,2002.
[6] 陈颙,陈凌. 分形几何学[M]. 北京 地震出版社,1998.
[7] 罗佳强. 齐波夫定律在济阳坳陷最终探明储量预测中的应用[J]. 石油实验地质,2003,25(3):318—320.
[8] 李景春. 试用齐波夫定律进行南金山金矿成矿预测[J]. 贵金属地质,1995,4(2):147—152.
[9] 周宏昆,丁宗强,雷祖志,等. 金属矿床大比例尺定量预测[M]. 北京 地质出版社,1993.
[10] 王新洲. 非线性模型参数估计理论与应用[M]. 武汉:武汉大学出版社,2002.

Fractal model of the reserve-size distribution of ore deposits and the mathematical analysis of its fractal dimension

WAN Li[1,2], DENG Jun[1], WANG Qing-fei[1], GAO Bang-fei[1]

(1. *School of Earth Sciences and Land Resources, China University of Geology, Beijing* 100083, *China*; 2. *School of Mathematics and Information Science, Guangzhou University, Guangzhou* 510405, *China*)

Abstract: It is main mean to establish mathematical models of distribution in the prediction of reserve-size of ore deposits, which is important in prospecting and evaluation of mineral resources. Based on statistical self-similarity of reserves of ore deposits and applying the methods of probability and statistics, a fractal model of reserve-size was established. The geological explanation of different fractal dimensions was presented via mathematical analyses. Two methods of estimation for fractal dimension, linear and non-linear least-squares regression estimation, were discussed.

Key words: reserves of ore deposits; distribution of reserve size; fractal model; fractal dimension