

分形分布函数及其在大型矿床中的应用

申 维 孙方勇

(中国地质大学国土资源与高新技术研究中心,北京,100083)

摘 要 分形概念应用在地球科学中来刻画地质质量和物体的自相似特征。研究表明,分形模型常常提供有力工具来刻画地质质量和物体的基本空间分布结构。许多地质现象具有尺度不变性,分形分布的特点要求大于某一尺度物体的数目,与物体大小之间存在着幂指数关系。本文论证了幂函数分布、帕累托分布、对数正态分布和齐波夫定律在一定的条件下具有分形不变性质,它们是分形模型的数学基础。基于分形模型,用求和方法确定中国山东省金地球化学元素异常值范围。等值线大于或等于金地球化学元素临界值(200×10^{-9})围成的异常面积包含了已知的大型、超大型金矿床。

关键词 分形分布 分形模型 地质异常值

Fractal Distribution and Its Application to Large - Sized Mineral Deposits

SHEN Wei SUN Fangyong

(Institute of High and New Techniques applied to Land Resources, China University of Geosciences, Beijing, 100083)

Abstract The concept of fractal has been used intensively in the field of earth science to characterize geological quantities and objects with self-similarity. It has been shown that fractal modeling often provides powerful tools for characterizing underlying spatial distribution structures of geological quantities and objects. Many geological phenomena are scale invariant. A fractal distribution requires that the number of objects larger than a specified size has a power-law dependence on the size. This paper shows that the power-function distribution, Pareto distribution, lognormal distribution and Zipf distribution have the fractal property under certain conditions and constitute the mathematical basis of fractal models. The summation method is developed on the basis of fractal models to determine thresholds of Au data in Shandong Province. The anomalous area enclosed by contours with contour values higher than or equal to the threshold (200×10^{-9}) contains the known large-sized and superlarge-sized gold mineral deposits.

Key words fractal distribution fractal model geological anomalous value

分形和多重分形产生于尺度独立性过程的自相似自然结果,并且应用在孔隙分布、网络(Gillepsie等,1993)、大小分布(Logdson,1995)和其他许多领域。许多地质现象具有尺度不变性,例如岩石碎片、断层、地震、火山喷发、矿藏和油井的频度和大小之间的分布。分形分布的特点要求大于某一尺度物体的数目,与物体大小之间存在着幂指数关系。在分形概念被提出之前很长一段时间,人们从经验上就认识到幂指数型的统计关系适用许多地质现象。

关于地质现象的统计分布,幂指数分布不是唯一,还有如正态分布和对数正态分布等其他种类。但是,幂指数分布是唯一不含特征分布长度的分布(Turcotte,1997)。这样,幂指数分布可以应用于那些具有标度不变性的地质现象。如果一定数目的事

件在统计上相互独立,那么,中心极限定理就提供了应用正态分布的基础。而标度不变性则提供了应用幂指数分形分布的基础。分形概念对连续分布的对象也适用。许多地球物理数据具有幂指数频谱。这包括有重力、地磁和地表地形。由于幂指数频谱由振幅和斜率两个量决定,因此利用这两个量可以进行数据集的结构分析。分析结构也可以用来作为对资料两点间进行内插值的依据。

1 分形分布

1.1 幂函数分布函数

1.1.1 一般的幂函数分布 如果随机变量 X 的密度函数为

$$f(x) = ak^{-a}x^{a-1} \quad a > 0 \quad 0 < x \leq k \quad (1)$$

则随机变量 X 服从幂函数分布(Johnson 等,1970)。其中 k 称为尺度参数, a 称为形状参数或指数。

由(1)式可求得随机变量 X 的分布函数,数学期望和方差,即

$$F(x) = P(X \leq x) = (x/k)^a \quad a > 0 \quad 0 < x \leq k \quad (2)$$

$$E(X) = ak(a+1)^{-1} \quad a > 0 \quad (3)$$

$$V(X) = ak^2(a+2)^{-1}(a+1)^{-2} \quad a > 0 \quad (4)$$

在高端截尾条件下 $0 < x \leq k' \leq k$, 推得下列等式

$$P(X \leq x | X \leq k') = (x/k')(k'/k)^{-a} = (x/k')^a \quad 0 < x \leq k' \leq k \quad (5)$$

上式表明,幂函数分布在高端截尾条件下具有分形性质(Shen 等,1998)。

$$P(X \leq x | X \leq k') = P(X \leq cx | X \leq ck') \quad 0 < x \leq k' \leq k \quad (6)$$

上式表明,幂函数分布具有尺度不变的特征。其中 k' 为尺度参数; c 为任意正的常数。

因此,幂函数分布是具有尺度不变的分形不变分布(在高截尾条件下)。

1.1.2 截尾的幂函数分布 如果随机变量 X 的密度函数为

$$f(x) = ak^{-a}xa^{-1}[1 - (k'/k)^a]^{-1} \quad a > 0, 0 < k' \leq x \leq k < \infty \quad (7)$$

则随机变量 X 服从截尾的幂函数分布。其中 k 与 k' 为尺度参数; a 为形状参数或指数。

由上式可求得随机变量 X 的分布函数

$$F(x) = P(X \leq x) = \frac{(x/k)^a - (k'/k)^a}{1 - (k'/k)^a} \quad a > 0 \quad 0 < k' \leq x \leq k < \infty \quad (8)$$

截尾的幂函数分布随机变量的数学期望和方差总是存在(当 $a > 0$ 时)。

在高端和低端截尾条件下, $0 < k' < k'_1 \leq x \leq k_1 < k < \infty$, 推得下列等式

$$P(X \leq x | k'_1 < X \leq k_1) = \frac{(x/k_1)^a - (k'_1/k_1)^a}{1 - (k'_1/k_1)^a} \quad a > 0, 0 < k'_1 \leq x \leq k_1 < \infty \quad (9)$$

上式表明,截尾的幂函数分布在高端和低端条件下具有分形性质。

$$P(X \leq x | k'_1 < X \leq k_1) = P(X \leq cx | ck'_1 < X \leq ck_1) \quad a > 0 \quad 0 < k'_1 \leq x \leq k_1 < \infty \quad (10)$$

上式表明,截尾的幂函数分布具有尺度不变的

特征,其中 k_1 和 k' 为尺度参数; c 为任意正的常数。

因此,截尾的幂函数分布是具有尺度不变的分形不变分布(在高端和低端截尾条件下)。

1.1.3 渐近的幂函数分布 如果随机变量 X 的密度函数为

$$f(x) \sim ak^{-a}xa^{-1} \quad a > 1, x \rightarrow 0 \quad (11)$$

则随机变量 X 服从渐近幂函数分布。其中 k 为尺度参数; a 为形状参数或指数。

渐近幂函数分布的函数为

$$F(x) = P(X \leq x) \sim (x/k)^a \quad a > 0, x \rightarrow 0 \quad (12)$$

上式可得出,当 $x \rightarrow 0$ 时,将数据对 $\{P(X \leq x_i)\}$ 和 $\{x_i\}$ ($i=1, 2, \dots, m$), 绘在双对数坐标纸上,它们渐近于一条直线上,直线的斜率为 a 。

1.2 帕累托分布函数

1.2.1 一般的帕累托分布 如果随机变量 X 的密度函数为

$$f(x) = ak^ax^{-a-1} \quad a > 0, x \geq k > 0 \quad (13)$$

则随机变量 X 服从帕累托分布(Johnson,1970),其中 k 称为尺度参数; a 称为形状参数或指数。

由(13)式可求得随机变量 X 的分布函数,数学期望和方差,即

$$F(x) = P(X \leq x) = 1 - (x/k)^{-a} \quad a > 0, x \geq k > 0 \quad (14)$$

$$E(X) = ak(a-1)^{-1} \quad a > 1 \quad (15)$$

$$V(X) = ak^2(a-2)^{-1}(a-1)^{-2} \quad a > 2 \quad (16)$$

曼德布罗特(Mandelbrot,1966)认为帕累托分布在低端截尾条件下具有尺度不变的分形性质,即:

$$P(X \geq x | X \geq k') = (x/k')^{-a} (k'/k)^a = (x/k')^{-a} \quad x \geq k' > k > 0 \quad (17)$$

上式表明,帕累托分布在低截尾条件下具有分形性质。

$$P(X \geq x | X \geq k') = P(X \geq cx | X \geq ck') \quad x \geq k' > k > 0 \quad (18)$$

式(18)表明,帕累托分布具有尺度不变的特征。其中 k' 为尺度参数; c 为任意正的常数。

因此,帕累托分布是具有尺度不变的分形不变分布(在低截尾条件下)。

1.2.2 截尾的帕累托分布 如果随机变量 X 的密度函数为

$$f(x) = ak^ax^{-a-1}[1 - (k'/k)^a]^{-1}$$

$$a > 0, 0 < k \leq x \leq k' < \infty \quad (19)$$

则随机变量 X 服从截尾的帕累托分布(Goldberg, 1969; Christopher 等, 1995)。其中 k 与 k' 称为尺度参数, a 称为形状参数或指数。

由(19)式可求得随机变量 X 的分布函数

$$F(x) = P(X \leq x) = 1 - \frac{(x/k)^{-a} - (k'/k)^{-a}}{1 - (k'/k)^{-a}}$$

$$a > 0, 0 < k \leq x \leq k' < \infty \quad (20)$$

截尾的帕累托随机变量的数学期望和方差总是存在(当 $a > 0$ 时)。

在高端和低端截尾条件下, $0 < k < k_1 \leq x \leq k'1 < k' < \infty$, 推得下列等式

$$P(X \geq x | k_1 < X \leq k'_1) = \frac{(x/k_1)^{-a} - (k'_1/k_1)^{-a}}{1 - (k'_1/k_1)^{-a}}$$

$$a > 0, 0 < k_1 \leq x \leq k'_1 < \infty \quad (21)$$

上式表明, 截尾的帕累托分布在高端和低端条件下具有分形性质。

$$P(X \geq x | k_1 < X \leq k'_1) = P(X \geq cx | ck_1 < X \leq ck'_1)$$

$$a > 0, 0 < k_1 \leq x \leq k'_1 < \infty \quad (22)$$

式(22)表明, 截尾的帕累托分布具有尺度不变的特征。其中 k_1 和 k' 为尺度参数; c 为任意正的常数。

因此, 截尾的帕累托分布是具有尺度不变的分形不变分布(在高端和低端截尾条件下)。

1.2.3 渐近的帕累托分布 如果随机变量 X 的密度函数为

$$f(x) \sim ak^ax^{-a-1} \quad a > 0, x \rightarrow \infty \quad (23)$$

则随机变量 X 服从渐近帕累托分布(Mandelbrot, 1982)。其中 k 称为尺度参数; a 称为形状参数或指数。

渐近帕累托分布的函数为

$$F(x) = P(X \leq x) \sim 1 - (x/k)^{-a} \quad a > 0, x \rightarrow \infty \quad (23)$$

Mandelbrot(1967)利用帕累托分布和它的渐近形式解释许多经验现象。Johnson 等(1970)讨论了帕累托分布的各种类型。

1.3 对数正态分布函数

Aitchison 等(1957)和 Crow 等(1988)详细讨论了对数正态分布。如果随机变量 $Y = \log X$ 服从正态分布, 则非负的随机变量 X 服从对数正态分布。

如果随机变量 X 的密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi\sigma x}} \exp \frac{-(\ln x - \mu)^2}{2\sigma^2} \quad x > 0 \quad (24)$$

则随机变量 X 服从对数正态分布(Law 等, 1991; Crovelli 等, 1997)。其中 μ 为尺度参数, $\sigma > 0$ 为形状参数。

由(24)式可求得随机变量 X 的分布函数, 数学期望和方差, 即

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \quad (25)$$

$$E(X) = e^{\mu + \sigma^2/2} \quad (26)$$

$$V(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1) \quad (27)$$

令 $X' = cX$, $c > 0$ 为常数, 则随机变量 X' 的密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi\sigma x}} \exp \frac{-(\ln x - \mu - \ln c)^2}{2\sigma^2} = \frac{1}{\sqrt{2\pi\sigma x}} \exp \frac{-(\ln x - \mu')^2}{2\sigma^2} \quad x > 0 \quad (28)$$

其中 $\mu' = \mu + \ln c$ 。

上式表明, 随机变量 X 和 $X' = cX$ 服从同一分布函数(对数正态分布), 且形状参数 σ 不变, 尺度参数 μ 改变为 $\mu + \ln c$, 即对数正态分布的随机变量在相乘一个正常数 $c > 0$ 的条件下, 分布形式不改变(形状参数 σ 不变, 尺度参数 μ 改变)(Mandelbrot, 1982)。因此, 对数正态分布在上述变换条件下, 具有分形性质。

1.4 齐波夫定律

如果随机变量的 Z 分布律为

$$P(Z = r) = ar^{-p} \quad (P > 1, r = 1, 2, \dots) \quad (29)$$

则随机变量 Z 服从齐波夫定律(Johnson 等, 1969)。

其中 $a = (\sum_{r=1}^{\infty} r^{-p})^{-1}$ 为常数, p 称为指数参数。

随机变量 Z 的分布函数, 数学期望和方差分别为

$$H(z) = P(Z \leq z) = \sum_{r \leq z} ar^{-p} \quad P > 1 \quad (30)$$

$$E(Z) = (\sum_{r=1}^{\infty} r^{-p+1}) / (\sum_{r=1}^{\infty} r^{-p}) \quad P > 1 \quad (31)$$

$$V(Z) = (\sum_{r=1}^{\infty} r^{-p+2}) / (\sum_{r=1}^{\infty} r^{-p}) - E^2(Z) \quad P > 1 \quad (32)$$

在低端截尾条件下, $r \geq k > 0$ 。推得下列等式

$$P(Z = r | Z \geq k) = (ar^{-P}) / (\sum_{r \geq k} ar^{-P}) \quad (33)$$

$$r \geq k > 0$$

此式说明齐波夫定律具有分形性质。

$$P(Z = r | Z \geq k) = P(Z = cr | Z \geq ck)$$

$$r \geq k > 0$$

其中 k 为尺度参数(正整数), c 为任意正的常数。此式说明齐波夫定律具有尺度不变性质。

因此,齐波夫定律是具有尺度不变的分形不变分布(在低截尾条件下)。

2 应用实例

2.1 分形模型

设分形模型

$$N(r) = Cr^{-D} \quad r > 0 \quad (34)$$

其中 r 表示特征尺度; $C > 0$ 为比例常数; $D > 0$ 为一般分维数(Shen 等,1998), $N(r) = N(\geq r)$ 表示尺度大于等于 r 的数目或和数。

许多地质现象具有标度不变的特征,如岩石碎片、断层、地震、火山喷发、矿藏和油井等(Christopher 等,1995)这些现象的频度和大小之间的分布具有尺度不变性,可以用幂指数分布加以描述。

为了求出其分形模型(34)中分维数 D ,将观测数据 $\{N(r_1), N(r_2), \dots, N(r_n)\}$ 和 (r_1, r_2, \dots, r_n) 绘在双对数坐标纸上,如果其点大致分布在一条直线上的话,分维数 D 可以利用直线的斜率求出,也就是说,将观测数据 $\{N(r_1), N(r_2), \dots, N(r_n)\}$ 和 (r_1, r_2, \dots, r_n) 代入(34)式,然后两边取对数(34)式化为一元线性回归模型

$$\text{Log}N(r) = -D\text{Log}r + \text{Log}C \quad (35)$$

用最小二乘法求出斜率 D 的估计量,即为分维数;如果其点大致分布在二段直线上时,可采用分段拟合,有的分界点清楚,有的则不清楚,为了提高分界点确定的客观性,在二个区间用最小二乘法回归时用了最优化方法确定分界点。其基本思想是,找出合适的分界点 r_{i0} ,使各区间拟合的直线与原始数据点之间的剩余平方和 $E_i (i = 1, 2)$ 在两个区间的总和

$$E = E_1 + E_2 = \sum_{i=1}^{i_0} [N(r_i) + D_1\text{Log}r_i - \text{Log}C_1]^2 + \sum_{i=i_0+1}^n [N(r_i) + D_2\text{Log}r_i - \text{Log}C_2]^2$$

为最小。其中 r_{i0} 是分界点, D_1 和 D_2 分别为相应区间的斜率即分维数。为了检验回归方程的显著

性,对每个回归方程都进行了相关系数检验及方差分析检验。

2.2 用求和方法确定地球化学元素异常值范围

设地球化学元素的值 $\{x_i\} i = 1, 2, \dots, n$ 。记

$$N(r) = \sum_{x_i \geq r}^n x_i \quad (36)$$

这里和式是对于所有满足 $x_i \geq r$ 的 i 求和,

$\text{Min}_{1 \leq i \leq n} x_i \leq r \leq \text{Max}_{1 \leq i \leq n} x_i$ 这样得到了数据 $(N(r_1), N(r_2), \dots, N(r_n))$ 和 (r_1, r_2, \dots, r_n) 。

山东省是中国著名的金矿化集中区和黄金产地,在此以金矿床的地球化学元素为例,用求和方法确定大型、超大型矿床的地球化学元素异常值范围。应用求和法(式36)对金元素 Au ($N = 21316$) 数据处理(表1)。

表1 用求和法(式36)求出 Au 元素数据的 $N(r)$

Table 1 $N(r)$ of Au data calculated by Eq.(36)

$r/10^{-9}$	$N(r)$	$r/10^{-9}$	$N(r)$
1	60246.63	100	16584.65
2	42225.09	120	15367.54
3	33972.22	150	14560.91
4	31415.72	200	13390.70
5	30200.14	250	11942.70
6	29336.08	300	11102.98
7	28693.53	350	8194.82
8	28234.95	400	6741.90
9	27080.11	450	5909.46
10	27444.68	500	5008.28
15	26078.95	600	3881.49
20	24886.95	650	3295.76
25	23892.90	700	2648.25
50	20691.04	800	1899.00

应用表1的数据,在 $\text{Log}r - \text{Log}N(r)$ 坐标中投点(图1),用最小二乘法拟合直线,得到相应的直线方程为

$$\text{Log}N(r) = -0.242938\text{Log}r + 4.68748$$

$$R^2 = 0.96399 \quad 1 \leq r < 200$$

$$\text{Log}N(r) = -1.41465\text{Log}r + 7.48322$$

$$R^2 = 0.95945 \quad 200 \leq r < 800$$

得到分维数 $D_1 = 0.242938, D_2 = 1.41465$,分界点 $r_{i0} = 200 (\times 10^{-9})$ 。以上方程均通过显著性检验。地球化学元素金的异常值范围为大于 200×10^{-9} 。

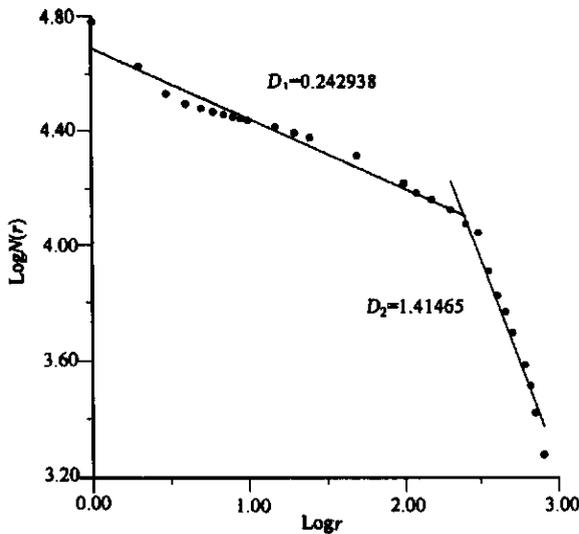


图1 $\text{Log}N(r) - \text{Log}r$ 图

Fig. 1 $\text{Log}N(r) - \text{Log}r$ plot for Au values from 21316 samples

图1显示 $N(r)$ 与 r 之间的关系。用最小二乘法拟合二段直线。 $r_{i0} = 200 \times 10^{-9}$ 表示金元素异常的临界点。用求和方法和其他方法(如面积方法)算出地球化学元素金的异常范围一致,且异常范围与大型、超大型金矿床的吻合程度很高,是寻找金矿的一种直接信息,对圈定靶区有着举足轻重的意义。引起分段分形的原因可能是大型、超大型矿床与中型、小型矿床成因不一。大型、超大型矿床是长期演化继承的地质产物,成矿系列有其有序性,其变量极值集合是大型、超大型矿床的控制因素。系列矿床共生和叠加成矿作用是大型、超大型矿床产生的必要条件。

3 结论

(1) 一般的幂函数分布在高端截尾条件下具有尺度不变的分形性质,截尾的幂函数分布在高端和低端条件下具有分形性质;一般的帕累托分布在低端截尾条件下具有尺度不变的分形性质,截尾的帕累托分布在高端和低端截尾条件下也具有分形性质。对数正态分布的随机变量在相乘一个正常数 $c > 0$ 的条件下,分布形式不改变(形状参数 σ 不变,尺度参数 μ 改变);齐波夫定律在低截尾条件下具

有分形性质。这些分布是分形模型的数学基础。

(2) 用求和方法确定中国山东省金地球化学元素异常值范围。等值线大于或等于金地球化学元素临界值(200×10^{-9})围成的异常面积包含了已知的大型、超大型金矿床,其结果与实际情况一致。该方法不仅适用于金元素和金矿床,而且还适用于其他元素和矿床,具有普遍的意义。

参考文献

- Aitchison J, Brown J A C. 1957. The Lognormal distribution. Cambridge: Cambridge University Press, 176.
- Christopher C B, Paul R L. 1995. Fractals in petroleum geology and earth process. New York: Plenum Press, 317.
- Crovelli R A, Schmoker J W, Balay R H. 1997. U S Department of the interior U S geological survey: Fractal lognormal percentage analysis of U S geological survey's 1995 national assessment of conventional oil and gas resources. Nonrenewable Resources 1(1): 43-51.
- Crow E L, Shimizu Kunio et al. 1988. Lognormal distributions-theory and applications. Mareel Dekker, New Inc., York and Basel, 387.
- Gillepsie P A, Howard C B, Walsh J J, Watterson J. 1993. Measurement and characterization of spatial distributions of fractures. Technophysics 226(1-4): 113-141.
- Goldberg G. 1967. The Pareto law and the pyramid distribution. Publication 505, Shell Development, Houston, Texas.
- John N L, Kotz S. 1969. Discrete distributions. John Wiley & Sons. Inc., New York, 328.
- John N L, Kotz S. 1970. Continuous univariate distributions-I. John Wiley & Sons. Inc., New York, 300.
- Law A M, Kelton W D. 1991. Simulation modeling and analysis (2nd ed.). McGraw-Hill, New York, 759.
- Logdson S D. 1995. Analysis of aggregate fractal dimensions and aggregate densities back-calculated from hydraulic conductivity. Soil Science Society of America Journal 59(5): 1216-1221.
- Mandelbrot B B. 1966. Forecasts of future prices, unbiased markets, and "martingale" models. J. Business, University of Chicago 39: 242-255.
- Mandelbrot B B. 1967. The variation of some other speculative prices. J. Business, University of Chicago 40: 393-413.
- Mandelbrot B B. 1982. The fractal geometry of nature. W H Freeman and Company, San Francisco, 460.
- Shen W, Zhao P D. 1998. Theory study of fractal statistical model and its application in geology. Scientia Geologica Sinica 33(2): 234-243.
- Turcotte D L. 1997. Fractals and Chaos in Geology and Geophysics (2nd ED.). Cambridge Univ. Press, Cambridge, 398.