

用逐步判别分析预测宁镇山脉中段铁矿

冯祖钧 朱慧娟 吴振荣 方长泉

前言

预测找矿有利地区的数学地质方法已相当多。不同方法适用于不同条件和目的,判别分析是其中之一。判别分析又分为若干种,各适应于不同情况和需要。本文是针对宁镇山脉中段地质及铁矿的研究现状,采用贝叶斯准则下的判别分析,在五万分之一比例尺的地质图上圈出找铁矿有利地段,而不是具体指出矿床位置与规模。由于资料及我们的水平所限,本文只是习作,不足为训。

研究地区的地质背景

本文研究地区是前江苏省重工业局1970年编制出版的1:5万《上党幅》和《汤山

幅》地质图的范围。宁镇山脉地质情况众所周知,兹不赘述(图1)。据生产单位意见,宁镇中段铁矿床以夕卡岩型为主。本文不谈成因类型问题,主要是依据本区已知的有利于找铁矿的地质条件进行数学地质计算。这一带现有一些小铁矿点,较大者只是个别的。据已知资料,与找铁矿有关的地质条件是:

1)下二叠统栖霞组、上石炭统船山组与中石炭统黄龙组等层位的灰岩,尤以栖霞组最重要。而中、下三叠统层位的灰岩较次要。其他层位更次要。

2)石英二长斑岩、石英二长岩及斑状石英二长岩较重要。而石英闪长斑岩较次要。

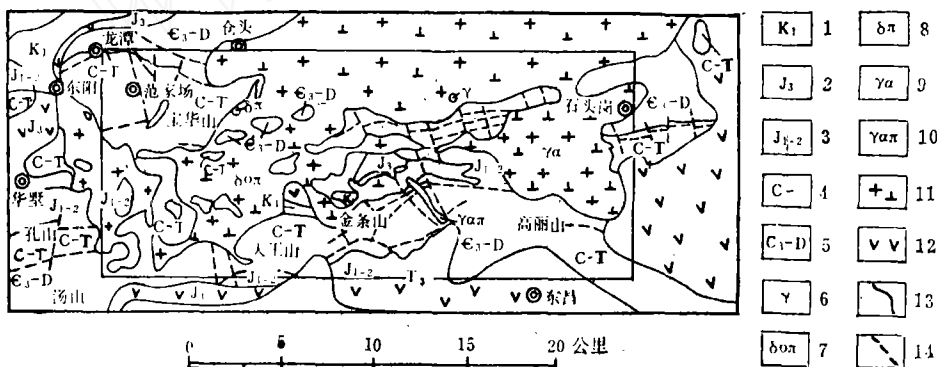


图1

1—下白垩统; 2—上侏罗统; 3—中、下侏罗统; 4—石炭—三叠系; 5—寒武系上统—泥盆系; 6—花岗岩; 7—石英闪长斑岩; 8—闪长岩; 9—石英二长岩; 10—石英二长斑岩; 11—中酸性侵入岩; 12—中酸性喷出岩; 13—地质界线; 14—断层

3)断裂构造,尤以深部的较重要。

此外,物探的重、磁异常具有很大指示性。

但现有这份1:5万地质图上,地层划分还较准确,而火成岩的分类则较粗略,断裂构造的表示精度很差。物探方面,除航磁等值线可供利用外,地面磁、重资料均连不成片,不能作区域性的数学地质计算。因之,本文就只用地质图上的地层岩性、火成岩、某些重要的接触界线等三种资料,配合

1:5万航磁等值线图,进行取值计算。

工作原理和方法

本文采用贝叶斯准则下的逐步判别分析进行计算。首先说明取值方法:

将两幅1:5万地质图(具有航磁等值线)划分成若干1×1公里的方格。划分时应严格与经纬度配合,但本图幅上掩盖太多,为了充分利用基岩出露面积,未按经纬度划分格子。只对有基岩分布的格子取值,

但一格中掩盖面积超过25%则不用。本图只取了156格。将每个方格划分成 1×1 毫米的400个小方格。每格中需要取值的地层或火成岩均数出其面积所占的小方格数，再换算成对400的百分比。如一格中有部分掩盖(小于25%)，则将剩下的基岩面积当成100%来计算各种岩性的百分比。接触界线按其1 $\times$ 1公里方格中的长度取值。因接触界线在1平方毫米的小格中可当成直线，故其长度就按所经过的小方格数取值。然后记下该接触线在方格中出现的次数。若断层线、褶皱轴线可以利用，则亦可按此方式取值。也可按1 $\times$ 1公里方格的中心点与断层线或褶皱轴线的垂直距离取值(仍以小方格数表示)。每格的中心点处的航磁等值线伽玛数即为其航磁取值。如采用地面物、化探等值线，则亦按此法取值。

被取值的各项指标就是为建立判别函数用的随机变量。本文用的随机变量共12个，依次为：

x_1 —中、下三叠统地层(T_{1-2})。灰岩与页岩。

x_2 —上二叠统地层(P_2)。页岩、砂岩、煤层。

x_3 —下二叠统，上、中石炭统地层(P_1-C_{2+3})。灰岩为主。

x_4 —下石炭统和上泥盆统地层(C_1-D_3)。砂岩、页岩，少量灰岩。

x_5 —志留系地层(S)。砂岩、页岩。

x_6 —石英闪长斑岩($\delta o \pi$)。

x_7 —石英二长斑岩、石英二长岩、斑状石英二长岩($\gamma a \pi$ 、 γa 、 $\pi \gamma a$)。

x_8 — P_1-C_{2+3} 与 C_1-D_3 接触界线长度。

x_9 — P_1-C_{2+3} 与 $\gamma a \pi$ 、 γa 、 $\pi \gamma a$ 的接触界线长度。

x_{10} —上述 x_8 的出现次数。

x_{11} —上述 x_9 的出现次数。

x_{12} —航磁值。

为了将全部1 $\times$ 1公里的方格区分成两组，即有矿组和无矿组，必须选取建立两组判别函数用的标准样，可以叫做控制格。按已知铁矿点选出16个有矿控制格。再按现有资料选出20个无矿控制格。同一组控制格不必相连，可分散在各处(图2)。

贝叶斯准则下的判别分析要求判别的母

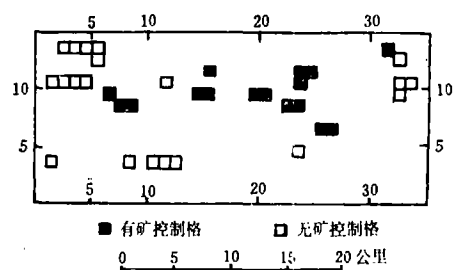


图 2

体应是多元正态分布。但上述随机变量在两组控制格内的取值几乎都不是正态分布，因而必须作变换处理。如果研究面积大，取样量大，控制格多，则可用对数变换、三角函数变换、方根变换等，使每个随机变量的实测数据经变换后，其偏度接近于零，峰度接近于三，即大体近似正态分布。但本文由于取样量较少，故只作了一些对数变换，未达到正态分布的要求。

两组判别函数建立后，就可计算出156个方格分别归属于有矿组或无矿组的后验概率值，亦即将某一格判为某一组的可靠程度的概率。最后再统一按归属于有矿组的后验概率值绘制等值线图，就能反映出找铁矿的有利地段了。

这种作法的理论根据如下：

关于判别分析本身这里不谈了，只说明如何把它应用于矿产预测。

设随机事件{找到铁矿}为 AB ；{铁矿生成}为 B ；{在铁矿存在的条件下找到铁矿}为 $A|B$ ，则按概率乘法定理有

$$P(AB) = P(B)P(A|B)$$

我们要求的是后验概率

$$P(A|B) = P(AB) / P(B)$$

这种关系也可用下式表示

$$P(x) = F(x)G(x)$$

式中 $P(x)$ 为找到 X 铁矿床的无条件概率； $F(x)$ 为 X 铁矿床生成的概率； $G(x)$ 为在 X 铁矿床存在的条件下而找到 X 铁矿床的概率。把 $G(x)$ 与 $F(x)$ 当作独立事件。显然， $F(x)$ 是研究问题的关键。但 $F(x)$ 不易知道。根据成矿理论，我们可以转而考虑下述的函数关系

$$V = Q(X_1, X_2, \dots, X_n)$$

$$P(V) = G(X_1, X_2, \dots, X_n, V)$$

式中 V 为铁矿床的某种计量； $P(V)$ 是 V

生成的概率, $X_1, X_2, \dots, X_n$ 是反映成矿地质作用的各项地质因素, 如地层、构造、火成岩等。

从理论上讲, 每项地质因素都是一个随机变量, 而且都有一定的数值范围。V 也是如此。因而它们的多维概率分布函数的一般式可写成

$$P_{V, X_1, X_2, \dots, X_n} = G(V, X_1, X_2, \dots, X_n)$$

从而可得 V 的条件概率分布函数为

$$P(V|X_1, X_2, \dots, X_n) = G(V, X_1, X_2, \dots, X_n) / G_{X_1, X_2, \dots, X_n}(V, X_1, X_2, \dots, X_n)$$

式中 $G_{X_1, X_2, \dots, X_n}(V, X_1, X_2, \dots, X_n)$ 为 V 的边际分布函数。

若已知函数 G 的表达式, 即可计算出 V 的条件概率。但函数 G 是未知的。

可以把每项地质因素 X_i 设想成一个线性组合, 即令

$$X_i = \sum_{j=1}^m a_{ij} Y_j \quad (i=1, 2, \dots, n)$$

式中 X_i 为理论上的地质变量; Y_j 是从地质图上实测的地质变量。对任一 X_i 来说, 某些 Y_j 的系数可以是零, 即可能有 $X_i = a_{ij} Y_j$ 。将每项实测变量通过适当变换, 使之接近正态分布。因而图上每个 1×1 公里的方格就成为一个个体, 其矢量符号为 $Y = (y_1, y_2, \dots, y_m)$ 。根据贝叶斯准则, 对于给定的个体 Y 来自母体 A_g 的条件概率 (后验概率) 为

$$P(g|Y) = q_g f_g(Y) / \sum_{i=1}^G g_i f_i(Y)$$

式中 G 为母体个数, $f_g(Y)$ 为母体 A_g 的概率密度函数; q_g 为先验概率。

若只有两个母体, 即有矿组和无矿组, 则按两组控制格建立两个判别函数, 就可求得计算每个 1×1 公里方格分组归属概率 (后验概率) 的表达式

$$P_1 = e^{f_1(Y_1, Y_2, \dots, Y_m)} /$$

$$\sum_{g=1}^2 e^{f_g(Y_1, Y_2, \dots, Y_m)}$$

$$(l=1, 2)$$

由此可见, 采用贝叶斯准则下的判别分

析来处理我们所设想的找铁矿的概率模式是合理的。

计算结果

先对两组控制格进行逐步判别。令引入变量和剔除变量的检验临界值相等, 即 $F_1 = F_2 = 2$ 。各步计算的主要结果列表如下 (表 1) :

表 1

计算步数	引入 (+) 或剔除 (-) 变量数	判别式中变量数	F	u	χ^2
1	+x ₇	1	35.16	0.492	23.79
2	+x ₁₁	2	4.21	0.436	27.39
3	+x ₈	3	4.24	0.385	31.01
4	+x ₁₂	4	4.17	0.339	34.57
5	+x ₆	5	2.21	0.316	36.27
6	+x ₉	6	4.34	0.275	40.02
7	+x ₅	7	3.58	0.244	43.05
8	-x ₁₁	6	1.45	0.256	42.19
9	+x ₁	7	2.23	0.237	43.85
10	+x ₁₁	8	2.80	0.215	46.09
11	-x ₈	7	1.42	0.226	45.29
12	+x ₃	8	3.15	0.203	47.86
13	-x ₈	7	0.26	0.205	48.37

表 1 中, F 为引入或剔除变量的检验统计量; u 为判别效果检验的威尔克斯 (Wilks) 统计量; χ^2 为 u 的近似检验统计量。由表 1 可见, 逐步判别计算结果引入 7 个变量, $u = 0.205$, $\chi^2 = 48.37$ 。查 χ^2 分布表, 得出

$$\chi_{0.01}^2[v(g-1)] = \chi_{0.01}^2[7] =$$

18.48 < 48.37 故两组差异显著, 判别效果良好。

求出两组判别式如下 (f_1 为有矿组, f_2 为无矿组) :

$$f_1 = 86.40x_1 + 117.56x_3 + 92.68x_5 + 65.50x_6 + 98.92x_7 + 352.67x_{11} + 127.63x_{12} - 420.17$$

$$f_2 = 75.77x_1 + 104.10x_3 + 81.27x_5 + 56.33x_6 + 82.77x_7 + 325.49x_{11} + 110.51x_{12} - 328.25$$

马哈拉诺比斯 (Mahalanobis) 距离: $D^2 = 14.85$;

马氏距离的 F 检验统计量: $F = 14.42$ 查 F 分布表, 得出

$$F_{\alpha}(v, N-g-v+1) = F_{\alpha}(7.28) =$$

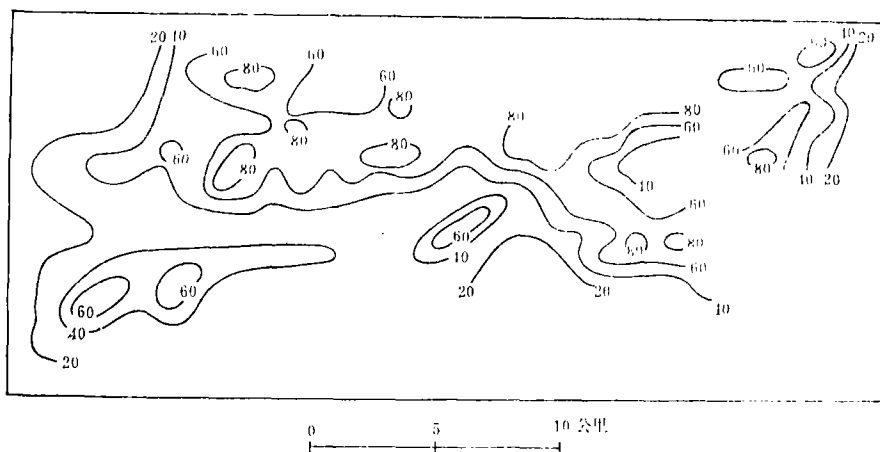


图 8

$$\begin{cases} 2.36, \alpha = 5\% \\ 3.36, \alpha = 1\% \end{cases}$$

都小于 $F = 14.42$ 很多, 两组差异大, 判别效果好。两组控制格的判别矩阵如下:

表 2

原分组 \ 计算分组	1	2	合 计
1	16	1	17
2	0	19	19
总 计	16	20	36

仅在无矿控制格中判错一格, 这说明上面建立的两组判别式与实际地质情况符合。

按贝叶斯定理, 计算出每个 1×1 公里方格归属于有矿组或无矿组的概率。再统一按归属有矿组的概率值绘制出 $1:5$ 万等值

线图 (图 3)。

由图可见, 80% 概率等值线圈定的地段自东向西有: ①石马庙, 戴家村东北面, 十里长山的西南端。②韦岗村以西十余里的上沟村一带。③香山、巢凤山、天王山等北东东一线。④桥头镇以南六、七里的谢家窑、朱家边一带。由此再向南五、六里的空青山、白庵一带。⑤范家塘向斜东端以南, 林家村一带。⑥由宝华山向西南方一带。

结 束 语

由于原始资料所限, 本文计算结果仅供参考。通过初步尝试, 我们认为用此法作矿产预测是可取的。若有精确的地质图及物、化探资料, 再有一些已知矿点的可靠数据, 将能达到较好的预测效果。

液动冲击回转钻进技术是八十年代金属矿床勘探的一种重要手段

苏联地质部把液动冲击回转钻进技术列为近十年重大技术成就之一。苏联用此种方法钻进的钻孔直径有 59、76、93 和 115 毫米。钻进岩层从 V 级到 VIII 级, 用硬质合金钻头钻进 V~VIII 级岩层 (部分 X 级岩层), 用金刚石钻头钻进 X~VIII 级岩层。最大钻进孔深达 1200 米。1976 年此种方法的钻进工作量达九十余万米, 计划 1980 年钻进 160 万米 (另一资料为 300 万米)。与钢粒钻进及金刚石钻进相比, 钻进每米岩层的成本降低 2~8 个卢布。近年来还研制了绳索取芯液动冲击器。

七十年代后期, 长春地质学院与有关单位研制成功了 $\phi 80$ 毫米的液压射流冲击回转钻具。有关兄弟部门研制成功了 $\phi 75$ 毫米的网式及液压射流式冲击回转钻具。目前地质系统有关单位正在研制 $\phi 54$ 毫米的网式冲击回转钻具、特种硬质合金钻头及用于深孔钻进的变量泥浆泵, 拟将液

动冲击回转钻进技术的研究, 列为部的重点科研项目。

根据苏联的试验资料, 冲击载荷比静压载荷能够更加有效地破碎脆性硬岩, 如花岗岩的静硬度为 70000 公斤/厘米², 而动硬度为 8000 公斤/厘米²。采用冲击回转钻进方法, 可以在较小的钻压和较低的转速下进行钻进, 可以减少钢材消耗, 减少孔斜, 减少孔内事故, 获得较高的机械钻进速度, 提高钻头寿命, 降低钻进成本。

在金属矿床勘探中, 遇到坚硬岩层的机会较多, 特别是遇到弱研磨性打滑地层, 如能采用小口径冲击回转方法, 可以获得较好的钻进效果。

(李常茂)

