



惩罚函数法在磁异常解释中的应用

左海燕

近年来,在物探资料推断解释过程中不断引入一些最优化反演算法^[1, 2]。无约束最优化算法往往因为多解性,效果一般都不太好。所以最近在地球物理反演理论方面,已逐渐注意研究在约束条件下的最优解问题。常用的数学工具是线性或非线性规划(带约束)方法^[3],本文着重讨论其中的惩罚函数法,并把它引入磁异常模型和实例的反演计算。计算结果表明,它确能减少反演解的多解性,比一般无约束方法能取得更有意义的结果。

方法原理概述

一般地球物理最优化反演的中心问题是,利用一定的数学方法去寻找一个“模型体”,使之与真实的地下某些地质结构的参数相符。换言之,即求一个 N 维参数向量 $X = (x_1, \dots, x_n)^T$,使计算的理论值向量 $f(X) = (f_1, \dots, f_m)$ 与观测值向量 $Y = (y_1, \dots, y_m)$ 尽可能符合。为此,一般可构造一个残差平方和比效函数:

$$\Phi(X) = [Y - f(X)]^T [Y - f(X)] \quad (1)$$

$\Phi(X)$ 为 $f(X)$ 与 Y 拟合程度的一个标志,也称为问题的目标函数。在一般最优化算法中,目的通常是求使 $\Phi(X)$ 取极小的 X 值。但在讨论约束最优化问题时,我们却要考虑如下形式的问题:

$$\min_{X \in S} \Phi(X)$$

$$S = \{X \mid h_i(X) \leq 0, i = 1, \dots, q\} \subset R^n \quad (2)$$

上式中 $X, \Phi(X)$ 的意义同前,集合 S 是 n 维空间 R^n 中的一个子集,称为可行域,它由约束方程组 $h_i(X)$ 所限制。一般,约束方程组的建立并不十分困难,也不复杂,因为在实际工作中总能了解到有关其他地球物理和地质方面的一些情况,据此即可建立相应的方程组。

惩罚函数法的基本思想是将约束问题(2)转化为如下形式的无约束问题:

$$\min_{X \in R^n} [\Phi(X) + \mu_k P(X)] \quad (3)$$

其中 $\{\mu_k\}$ 为一单调递增常数序列,即 $\mu_1, \mu_2, \dots, \mu_k, \dots, \lim_{k \rightarrow \infty} \mu_k = \infty$ 。 μ_k 也叫惩罚因子,

$\mu_k P(X)$ 称作惩罚项。 $P(X)$ 为惩罚函数,应满足下列条件:

- (1) $P(X)$ 是 R^n 上的连续函数;
- (2) $P(X) > 0$, 当 $X \notin S$ 时;
- (3) $P(X) = 0$, 当 $X \in S$ 时。

惩罚函数通常由约束方程 $h_i(X)$ 构成,其形式如下:

$$P(X) = \sum_{i=1}^q [\max\{0, h_i(X)\}]^2 \quad (5)$$

由此我们可将(3)式定义为增广的目标函数 $F(X, \mu_k)$,则有:

$$\min_{X \in R^n} F(X, \mu_k) = \min_{X \in R^n} [\Phi(X) + \mu_k \sum_{i=1}^q (\max\{0, h_i(X)\})^2] \quad (6)$$

利用惩罚函数将约束问题化为无约束问题后,即可用一般的无约束最优化算法解算。至于惩罚函数法的理论和收敛性证明等,由于篇幅所限,本文不作评述,请参看文献[3]·[4]。以适用物探解释为原则,我们首先引用收敛较快的阻尼最小二乘法。为了将惩罚项引入该方法的正则方程中进行运算,我们作了如下修改:设约束方程 $h_i(X)$ 为线性的,并可写成方程组形式:

$$h_i(X) = BX - C \leq 0 \quad (7)$$

式中 B 为 q 行 n 列矩阵, C 与 0 为 q 维列向量。为便于建立矩阵形式的公式,引入:

$$w_i = \begin{cases} 1, & \text{当 } h_i(X) > 0 \\ 0, & \text{当 } h_i(X) \leq 0 \end{cases}$$

$$W = \begin{bmatrix} w_1 & & 0 \\ & w_2 & \\ 0 & & \dots & w_q \end{bmatrix} \quad (8)$$

因此(5)式可写为:

$$\begin{aligned} \mu P(X) &= \mu \sum_{i=1}^q [w_i h_i(X)]^2 \\ &= \mu [BX - C]^T W [BX - C] \end{aligned} \quad (9)$$

若将 $X = X^0 + \delta$ (δ 为参数增量, X^0 为初始值) 代入上式,则有:

$$\mu P(X^0 + \delta) = \mu \{ [BX^0 - C]^T W [BX^0 - C] + 2(B\delta)^T W [BX^0 - C] + \delta^T B^T W B \delta \} \quad (10)$$

对上式中的 δ 求导数,并令该导数为零:

$$\begin{aligned} \frac{\partial \mu P(X^0 + \delta)}{\partial \delta} &= \mu \{ 2 B^T W [BX^0 - C] + 2 B^T W B \delta \} \\ &= 0 \end{aligned} \quad (11)$$

将 $X = X^0 + \delta$ 代入(1)式,并对 $\Phi(X^0 + \delta)$ 作泰勒展开,取二级近似,连同(10)式代入增广目标函数(6)得:

$$\begin{aligned} F(X^0 + \delta, \mu) &= \Phi(X^0 + \delta) + \mu P(X^0 + \delta) \\ &= [Y - f(X^0)]^T [Y - f(X^0)] - 2[Y - f(X^0)]^T G \delta + \delta^T G^T G \delta \\ &\quad + \mu \{ [BX^0 - C]^T W [BX^0 - C] + 2(B\delta)^T W [BX^0 - C] + \\ &\quad + \delta^T B^T W B \delta \} \end{aligned} \quad (12)$$

上式中 G 为雅可比微分矩阵:

$$G = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & & \frac{\partial f_m}{\partial x_n} \end{bmatrix}_{x=X^0} \quad (13)$$

对(12)式中的 δ 求导数,并令导数为零,利用(11)式得:

$$\begin{aligned} \frac{\partial F(X^0 + \delta, \mu)}{\partial \delta} &= \frac{\partial \Phi}{\partial \delta} + \mu \frac{\partial P}{\partial \delta} \\ &= -G^T [Y - f(X^0)] + G^T G \delta + \mu \{ B^T W [BX^0 - C] + B^T W B \delta \} \\ &= 0 \end{aligned} \quad (14)$$

经整理后,并由(9)式知(下式中 ∇ 为梯度符号):

$$\begin{aligned} \nabla P(X^0) &= 2[B^T W (BX^0 - C)] \\ \nabla \Phi(X^0) &= -2G^T [Y - f(X^0)] \end{aligned} \quad (15)$$

且令

$$A = [G^T G + \mu B^T W B]$$

$$g = -\frac{1}{2} \nabla [\Phi(X^0) - \mu P(X^0)] \quad (16)$$

将上式代入 (14) 最后得到修改后的正则方程:

$$A\delta = g \quad (17)$$

将 (17) 式加上阻尼因子, 按阻尼最小二乘法一般步骤计算即可。带有惩罚函数的阻尼最小二乘法反演 Fortran 程序已在 M-160 II 机上调试通过 (地质部计算中心)。

下面给出惩罚函数法简单的计算步骤:

(i) 取 $\mu_1 > 0$, 给定初始值 $X^0 \in R^n$ 和允许误差 ϵ ;

(ii) 设已利用某种无约束最优化方法得到式 (6) 对应于 μ_k 的近似极小点

$$X^*(\mu_k) \in S;$$

(iii) 以 $X^*(\mu_k)$ 为初始点, 取 $\mu_{k+1} > \mu_k$ (一般可令 $\mu_{k+1} = t\mu_k$, t 为一固定正常数)。继续利用无约束最优化方法求解问题: $\min_{x \in R^n} F(X, \mu_{k+1})$, 得到相应的极小点 $X^*(\mu_{k+1})$;

(iv) 若对一切 i ($i = 1, 2, \dots, q$) 均能满足:

$$h_i[X^*(\mu_{k+1})] < \epsilon \quad (18)$$

则可取 $X^*(\mu_{k+1})$ 为满足精度要求 (6) 式的极小点, 否则以 μ_{k-2} 代替 μ_{k-1} , 并以 $X^*(\mu_{k-1})$ 为新的初始近似值, 返回去作第 (iii) 步, 继续迭代计算。

惩罚函数法模型与实例计算

1. 模型实验 在算法的模型实验中, 为了应用上的广泛性, 我们所选的模型是磁法中二维斜磁化倾斜板状体, 其垂直场强 Z_a 的正演公式如下:

$$Z_a(x_k) = 2J \sin \alpha \left\{ \frac{1}{2} \sin \gamma \ln \frac{[(h-l \cdot \sin \alpha)^2 + (x_k - x_0 - b + l \cdot \cos \alpha)^2]}{[(h-l \cdot \sin \alpha)^2 + (x_k - x_0 + b + l \cdot \cos \alpha)^2]} + \frac{[(h+l \cdot \sin \alpha)^2 + (x_k - x_0 + b - l \cdot \cos \alpha)^2]}{[(h+l \cdot \sin \alpha)^2 + (x_k - x_0 - b - l \cdot \cos \alpha)^2]} + \cos \gamma \left[\operatorname{tg}^{-1} \frac{2b(h-l \cdot \sin \alpha)}{(h-l \cdot \sin \alpha)^2 + (x_k - x_0 + l \cdot \cos \alpha)^2 - b^2} - \operatorname{tg}^{-1} \frac{2b(h+l \cdot \sin \alpha)}{(h+l \cdot \sin \alpha)^2 + (x_k - x_0 - l \cdot \cos \alpha)^2 - b^2} \right] \right\} \quad (19)$$

上式中有 7 个参变量 (图 1):

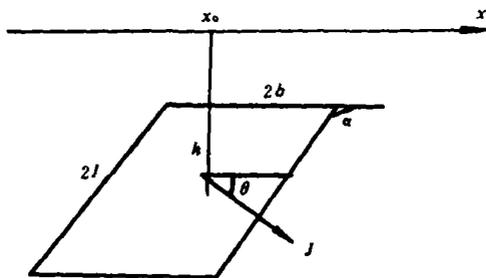


图 1

x_0 —模型中心横坐标 (米);

h —模型中心纵坐标 (米);

$2b$ —模型宽度 (米);

$2l$ —模型延伸长度 (米);

α —模型倾角 (度);

θ —磁化方向角 (度);

J —磁化强度值 (10^{-6} CGSM)

计算中使用前述加惩罚项的阻尼最小二乘法和单纯型法。理论模型参数和初始值的给出是经过选择的, 即

选用那些曾用无约束最优化方法计算过,但收敛不好或解的效果不好的参数,再用惩罚函数法加以计算后,进行对比,计算结果可参看下列图表(注意:约束方程是用人为因素大致给定的,无特殊限制)。

表1~表6给出图2~图6中模型实验各种数据,其中包括理论模型参数值,初始假设模型值反演迭代计算结果以及拟合的目标函数 Φ , F 值等(表中参量单位见前面说明)。

图2和图3对应于表2、表3,是用一般无约束方法计算的结果^[5]。从图、表可看出,尽管 Φ 值很小,曲线拟合良好,但反演计算模型体已面目皆非,无法辨认。两图使用同一理论模型,但初始假设不同,用初值I计算的结果, h 为负(表明计算得到的模型体是在地表之上),用初值II,迭代90次以上,计算结果甚差,以至无法图示。

表4对应于图4,其理论模型和初值的假设值均与表2同,不过这里是用惩罚函数阻尼最小二乘法计算的,30次迭代后,不但曲线拟合好,且计算结果与理论模型几乎重合。约束方程为: $x_0: 200, h < 70, 2l - 4b < 20, a < 120, \theta < 100, J < 1600$ (20)

理论模型体参量值

表1

参量	x_0	h	$2b$	$2l$	a	θ	J
模型1	160	60	20	50	120	90	1500
模型2	180	23	10	15	90	45	1414

无约束最优反演(理论模型1)

表2

参量 迭代次数	x_0	h	$2b$	$2l$	a	θ	J	Φ
初值I	150	70	30	60	60.5	45	1000	$3.15 \times (10^6)$
1次迭代	172.1	75.4	37.4	53.0	10.5	74.5	1047	$1.76 \times (10^6)$
10次迭代	160.9	-56.1	20.0	40.0	43.5	90.5	2200	$9.45 \times (10^2)$

无约束最优化反演(理论模型1)

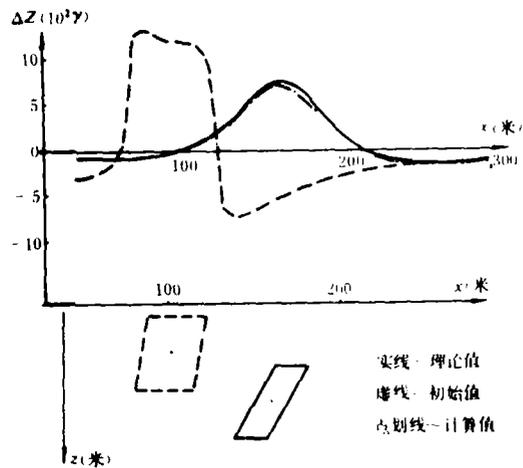
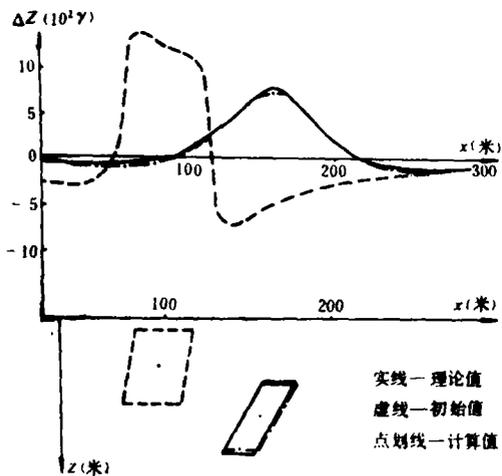
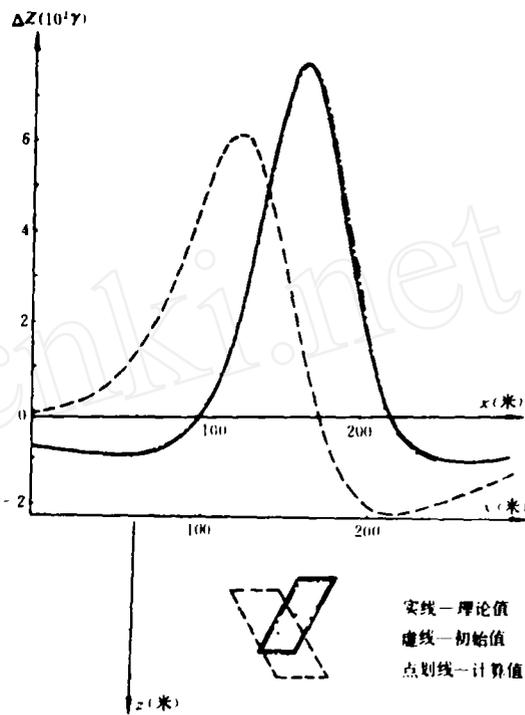
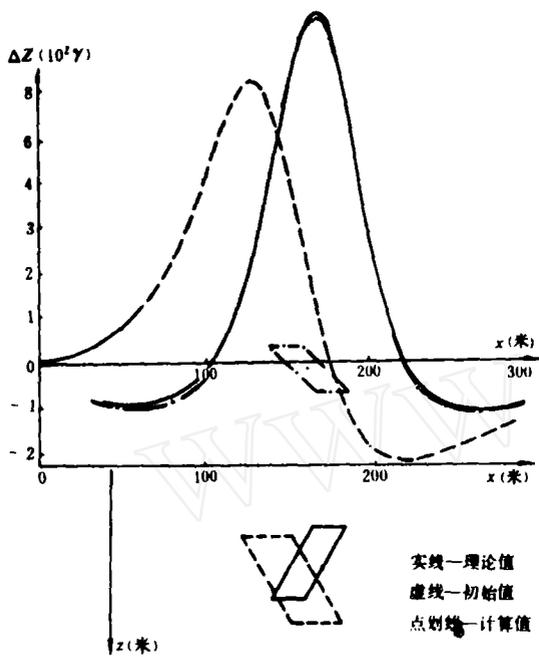
表3

参量 迭代次数	x_0	h	$2b$	$2l$	a	θ	J	Φ
初值II	100	30	35	45	100	80	1000	3.5×10^7
1次迭代	91.4	28.8	34.8	40.7	122.3	101.2	54.9	2.5×10^6
50次迭代	166.0	-55.8	-17.5	-15.9	-5770	-9625	-23632	2.9×10^3
90次迭代	166.2	-56.1	-15.4	-13.7	-5772	-9625	-27806	2.7×10^3

惩罚函数法反演(理论模型1)

表4

参量 迭代次数	x_0	h	$2b$	$2l$	a	θ	J	$F(\Phi)$
初值I	150	70	30	60	60	45	1000	$3.15 (10^6)$
1次迭代	169.2	78.4	37.9	51.8	13.2	76.6	1073.2	$1.6 (10^6)$
10次迭代	165.5	57.7	26.4	29.9	98.2	85.6	1606.1	$1.6 (10^3)$
30次迭代	160.0	60.0	20.0	50.0	120	90.0	1500	$1.7 (10^{-10})$



逐初函数算法 (理论模型)

表 5

参量	x_0	h	$2b$	$2l$	a	θ	J	$F(\Phi)$
迭代次数								
初值 Φ	100	30	35	45	100	80	1000	$3.5 (10^4)$
1次迭代	83.5	28	30.9	42	151.9	110.4	275.9	$2.9 (10^6)$
10次迭代	160.4	60.5	16.6	49.5	111.4	89.5	1600.1	$7.9 (10^3)$
15次迭代	161.3	61.9	17.5	51.2	112.1	88.9	1604.6	$2.6 (10^2)$

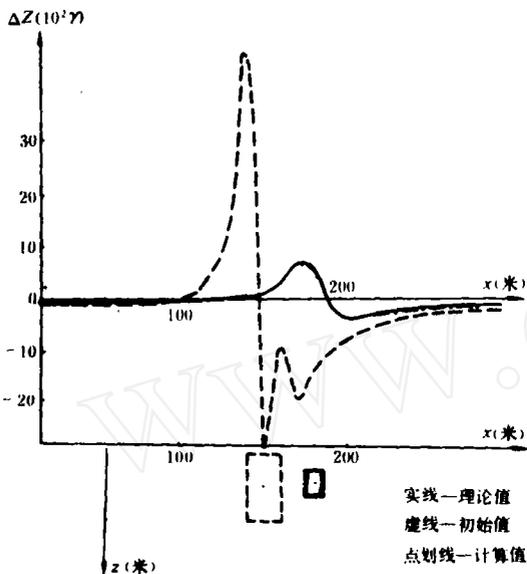


图 6

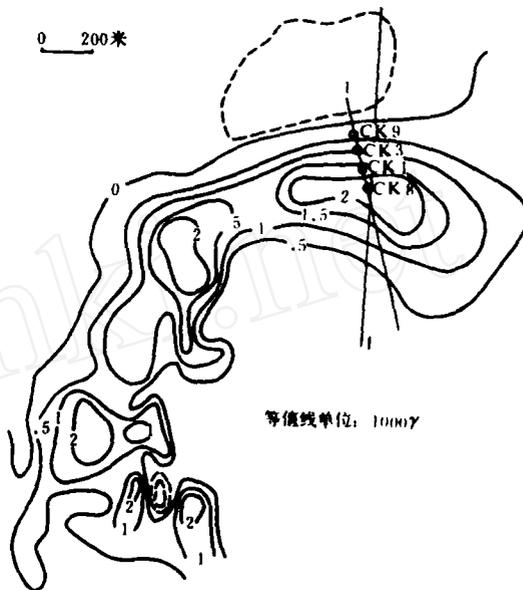


图 7

惩罚函数法反演 (理论模型 2)

表 6

参数	x_0	h	$2b$	$2l$	a	θ	J	$F(\Phi)$
迭代次数								
初值 III	150.0	25	20	40	90	45	1414	$4.3 (10^7)$
1次迭代	147.9	28	20.6	46.3	93.7	54.7	160.3	$2.1 (10^6)$
10次迭代	181.0	21.8	14.3	10.9	40.4	43.2	1772.2	$3.2 (10^4)$
20次迭代	180.0	22.9	9.7	14.7	89.8	45.0	1496.5	$2.7 (10^2)$

表 5 为图 5 的对应值, 理论模型与初始值均与表 3 同, 也是用惩罚函数法计算的。与图 3 比较, 显然带约束的方法效果好。约束方程给出如下: $x_0 \leq 200$, $h \leq 100$, $2l - 4b \leq 20$, $a \leq 150$, $\theta \leq 100$, $J \leq 1600$ (21)

表 6 对应于图 6, 所设计的第二个理论模型为直立板, 模型实验结果示于图 6。理论参数为表 1 中模型 2 的值, 约束方程如下设立:

$$x_0 \leq 200, h \leq 40, 2l - 4b \leq 10$$

$$a \leq 90, \theta \leq 70, J \leq 1500 \quad (22)$$

从图、表可看出, 尽管异常曲线形态复杂, 用带约束的阻尼最小二乘法取得的效果仍然很好, 这是用一般方法往往难以收敛的。

对带惩罚项的单纯型法也曾做过类似的模型试验, 效果也较好。当然还缺乏对各种带惩罚函数的最优化算法逐一做详尽的对比实验, 原因在于要消耗大量计算机运算时间, 另外就对比而言, 确定一个客观而广泛适用的标准也是困难的。我们只能在有限范围内做一粗略的比较。我们还对以上反演过程中的约束方程进行过不同改变, 发现除迭代次数和计算时间略有不同外, 计算结果几乎都一致是令人满意的, 这对该方程的可行性提供了一定的保证。

2. 实例 我们利用上述带惩罚函数的阻尼最小二乘法对某地区崇义铁矿西部矿体异常进行了反

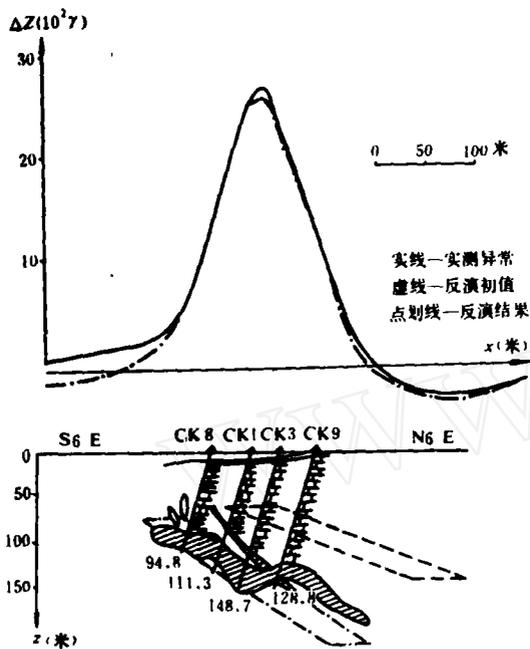


图 8

以及用惩罚函数法反演的结果一并绘在 I 号异常地质勘探剖面的草图上, 以便进行比较。从图中看到尽管所选初值并不理想, 但计算结果与实际矿体位置大致相符, 异常曲线与理论计算曲线中段拟合较好, 从而说明惩罚函数法处理实际资料是合适的。

上述模型实验与计算实例均在 6912-5 机和 M-160 II 机上实现, 在各种情况下所耗时间较短, 收敛很快。

结束语

为了减少物探资料推断解释的多解性, 除了应进一步提高数据采样的可靠程度和计算精度, 增强有益信息的分辨能力外, 主要的办法是对反演过程进行约束, 即尽可能地利用各种已知地球物理和地质资料 (也称先验信息) 对解的范围加以适当限制, 才可能取得理想的结果。根据对磁异常模型计算实验和实际资料处理, 我们认为用惩罚函数法解决带约束的最优化问题是比较有效的, 而且方法简便实用, 可望在其他地球物理资料反演中得到推广使用。

工作中得到朱介寿、周熙襄、钟本善等同志的支持, 在此表示感谢。

参考文献

- [1] 周熙襄等, 地球物理反演问题中的最优化方法, 地质出版社, 1980
- [2] 南京大学数学系计算数学专业, 最优化方法, 科学出版社, 1978
- [3] M. Arriel (李元薰等译), 非线性规划——分析与方法, 上海科技出版社, 1979
- [4] 王德人, 非线性方程组解法与最优化方法, 人民教育出版社, 1979
- [5] 成都地质学院物探系金属物探电算组, 金属矿物探电算文集, 地质出版社, 1979