

· 煤田物探 ·

区域物探中 KL 变换的数字滤波法

夏克文 (西安石油勘探仪器总厂测井研究所 西安 710061)

摘要 提出了在区域物探资料处理中的 KL 变换法, 对区域物探数据进行 KL 变换、提取特征和重构数据, 可以将随机噪声降低到最低限度。人工资料和实际资料表明: 这种数字滤波方法效果很好, 大大提高了物探数据的信噪比。

关键词 数字滤波; KL 变换法, 地球物理勘探

中国图书资料分类法分类号 P631.14; P631.24

在区域物探(重、磁、电)资料处理中, 剔除随机噪声是常规处理步骤之一。最简单的方法是徒手圆滑; 常用的还有最小二乘滤波法等, 它们都能有效地剔除局部随机噪声。本文从全局(整个测区范围)出发, 应用在地震、声波全波列测井资料处理中用来提取信号特征的 KL (Karhuen—Loeve) 变换数字滤波法^[1]进行位场数据去噪处理。

1 方法原理

KL 变换也即主成分分析, 对信号波形数据进行波形特征提取并重构波形, 因此又称为特征向量变换, 是一种正交变换。它与 Fourier 变换、Walsh 变换相似, 可用矩阵表示为: $\mathbf{Y} = \mathbf{T}\mathbf{X}$, 式中, \mathbf{X} 是要变换的矢量; \mathbf{T} 是变换矩阵; \mathbf{Y} 是变换后的矢量。KL 变换有着两个最大特点: 一是唯一产生不相关系数的变换, 有利于图象数据压缩; 二是它的均方差最小, 是一种最佳变换。因此, 它在正交变换中占有独特的位置, 是一种理想的变换。

在区域物探测量中所观测的数据是信号与噪声的叠加, 而噪声是由各种干扰和误差造成的。为将噪声压低到最低限度, 我们从整个测区范围的角度采用 KL 变换进行处理。

假定实测区域物探数据可以构成一个 $N \times P$ 阶矩阵 \mathbf{X} , 例如, 数据由 P 条剖面、每

条剖面有 N 个测点组成, 或者是由 N 个测点、每个测点有 P 个频率或时间取样道的数据组成。若将每道波形数据看成为一个离散随机信号, 则 \mathbf{X} 可表示为一个 N 维随机向量: $\mathbf{X} = [x_1, x_2, \dots, x_N]^T$, 其中, $x_i = [x_{i1}, x_{i2}, \dots, x_{ip}]$ ($i = 1, 2, \dots, N$)。其均值向量为: $\mathbf{mX} = \mathbf{E}(\mathbf{X}) = [E(x_1), E(x_2), \dots, E(x_N)]^T$, \mathbf{X} 的协方差矩阵记为: $\mathbf{CX} = \mathbf{E}[(\mathbf{X} - \mathbf{mX})(\mathbf{X} - \mathbf{mX})^T]$ 。 \mathbf{CX} 为实对称矩阵, 将其特征值按其递减顺序排列为: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ 。与特征值对应的特征向量 \mathbf{a}_i 满足关系: $\mathbf{CXa}_i = \lambda_i \mathbf{a}_i$ ($i = 1, 2, \dots, N$)。用特征向量组 $\{\mathbf{a}_i\}$ 组成正交变换矩阵为: $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N]$, $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N]^T$ 为任一向量 \mathbf{X} 展成 $\mathbf{X} = \sum_{i=1}^N \mathbf{Y}_i \mathbf{a}_i$ 的展开系数矢量。我们称 \mathbf{Y} 为 \mathbf{X} 的 KL 变换, 或 \mathbf{Y}_i ($i = 1, 2, \dots, N$) 为 \mathbf{X} 的主分量。由于 \mathbf{A} 为正交归一化矩阵, 则 \mathbf{X} 的准确表达式为:

$$\mathbf{X} = \mathbf{AY} \quad (1)$$

若将 \mathbf{X} 分为主要信号 \mathbf{X}_s 和非主要信号 \mathbf{X}_N , 则:

$$\mathbf{X} = \mathbf{X}_s + \mathbf{X}_N \quad (2)$$

利用 KL 变换, 可将 \mathbf{X}_s 与 \mathbf{X}_N 分离开。其出发点基于如下原理: 主要信号通常具有较大的能量, 从信号的某个特征出发, 它们在各道之间具有较强的相干性, KL 变换后, 这些相干

性能量主要集中在 X 的协方差矩阵 CX 的前 N 个最大特征值相对应的主分量上; 而那些非主要信号具有的能量较小, 各道间的相干性也很弱, 它们与那些具有较小特征值的主分量相对应。由于各主分量之间互不相关, 故可通过与 CX 较大特征值相对应的主分量来构成主要信号, 而用与较小特征值对应的主分量来构成非主要信号, 从而达到信噪分离的目的。若主要信号 X_s 由前 M 个主分量构成, 则由(1)、(2)式可得:

$$X_s = \begin{bmatrix} \mathbf{a}_{11} \mathbf{a}_{12} \dots \mathbf{a}_{1m} \\ \mathbf{a}_{21} \mathbf{a}_{22} \dots \mathbf{a}_{2m} \\ \dots \\ \mathbf{a}_{N1} \mathbf{a}_{N2} \dots \mathbf{a}_{NM} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_{11} \mathbf{Y}_{12} \dots \mathbf{Y}_{1P} \\ \mathbf{Y}_{21} \mathbf{Y}_{22} \dots \mathbf{Y}_{2P} \\ \dots \\ \mathbf{Y}_{M1} \mathbf{Y}_{M2} \dots \mathbf{Y}_{MP} \end{bmatrix} \quad (3)$$

这样, 重构的信号去掉了 X_N 。重构能量占原能量的百分数 $\eta(M)$ 为:

$$\eta(M) = \left(\sum_{i=1}^M \lambda_i / \sum_{j=1}^N \lambda_j \right) \times 100\% \quad (4)$$

均方差为:

$$\begin{aligned} \bar{\sigma}(M) &= \left(\frac{E\{\|X - X_s\|^2\}}{N \times P} \right)^{\frac{1}{2}} \\ &= \left(\frac{1}{N \times P} \sum_{i=M+1}^N \lambda_i \right)^{\frac{1}{2}} \end{aligned} \quad (5)$$

从理论上讲, 均方差 $\bar{\sigma}(M)$ 应接近随机噪声的均方差, 即 $\bar{\sigma}(M) \leq \epsilon$, 这是重构资料剔除噪声的必要条件。因此, 提取特征选取 M 的原则可从理论上得到, 即当 $\bar{\sigma}(M) \leq \epsilon$ (实测资料的均方差) 时的 M 为所求。我们可利用(4)式, 令 $M=1$, 计算 $\bar{\sigma}(1)$, 然后检验 $\bar{\sigma}(1) \leq \epsilon$? 如果 $\bar{\sigma}(1)$ 不大于 ϵ , 则选取 $M=1$ 。否则令 $M=2$ 。重复上述步骤, 直至 $\bar{\sigma}(M) \leq \epsilon$, 则可选定 M 。另外我们可根据 $\eta(M)$ 曲线形态判定。 $\eta(M)$ 曲线揭示了信噪变化规律。从下一节理论模型图 3 可知, 无噪声时曲线随 M 递增迅速使 $\eta(M)$ 等于 1; 有噪声时曲线随 M 递增不是迅速使 $\eta(M)$ 为 1 而是缓缓地趋于 1, 曲线接近一条水平直线。此水平直线是为随机

噪声所致, 故选取 $\eta(M)$ 曲线开始变为直线时的位置所处的 M 即为所求。

2 人工资料处理

模型 I: 我们以重力数据处理为例进行讨论。首先用若干理论模型通过计算叠加成面积上的理论重力数据, 并在这个理论数据中加入满足高斯分布的随机噪声, 构成带有噪声的数据, 然后对这样的数据进行 KL 变换提取特征和重构数据。

图 1 是由若干个球体在地面产生的异常等值线图, 其中剖面测线为 x 方向, 每条剖面测线上的采样点 $N=40$, 剖面测线数 $P=40$, 异常的极大值为 150 mgal。在重力测量中, 干扰因素较多, 尽管在接收系统中已将噪声压到最小限度, 但在测量记录上随机噪声仍明显可见。为模拟实际噪声特征, 我们用

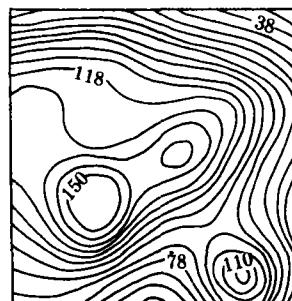


图 1 由若干个球体所构成的重力异常图
(等值线间隔为 8 mgal)

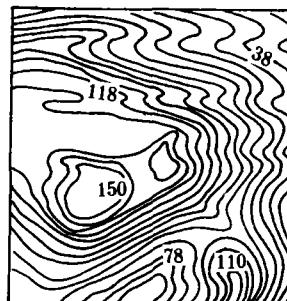


图 2 带有噪声数据的平面图

Monte-Carlo 法在微机上计算出服从高斯分布的随机序列,然后由此构造出具有实际意义的离散噪声序列,则得到 $N \times P$ 个均值近似为零,均方差为 9 mgal 的随机噪声(相当于异常极大值的 6%)叠加在理论数据上。

图 2 是用加入噪声后的数据绘成的平面图,显见其等值线变得零乱。对图 2 数据进行 KL 变换,求出主分量所对应的特征值,再由(4)式绘出 $\eta(M)$ 曲线(图 3)。由图 3 可见, M 从 4 开始曲线变成一条水平直线。根据均方差分析, $\bar{\sigma}(1) = 39.9$, $\bar{\sigma}(2) = 18.6$, 而 $\bar{\sigma}(3) = 7.9$, 小于随机噪声的均方差 $\epsilon = 9$ mgal。因此我们提取前三个特征,应用公式(3)重构数据,得出消除噪声到最低限度的平面等值线图 4。对比图 1 可知,异常形态吻合得很好,也即图 2 去噪后基本复原。

模型 I: 由于重磁异常测量不可能保证在平面上进行,因而在曲面上测得的数据还需“曲化平”。为不使曲面上的随机噪声(测量

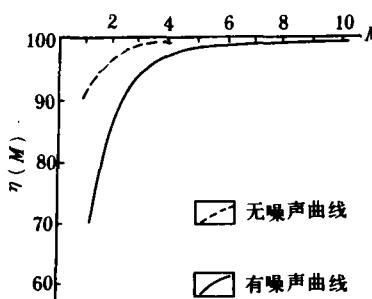


图 3 $\eta(M)-M$ 关系曲线图

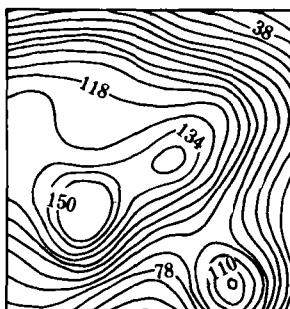
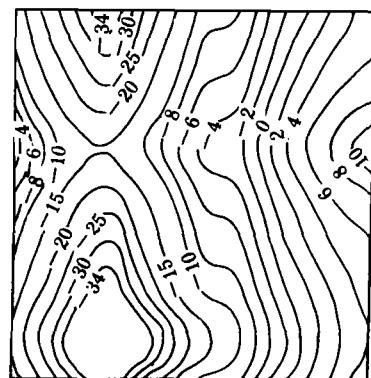


图 4 重构数据的平面图

表 1 垂直直角棱柱体的尺寸及磁参数

形体四角坐标(m)						磁化强度 ($\times 10^{-3}$ A/m)	磁化 倾角 I	磁化 偏角 D
x_1	x_2	y_1	y_2	z_1	z_2			
45.0	75.0	45.0	75.0	30.0	90.0	17 000.0	0°	0°



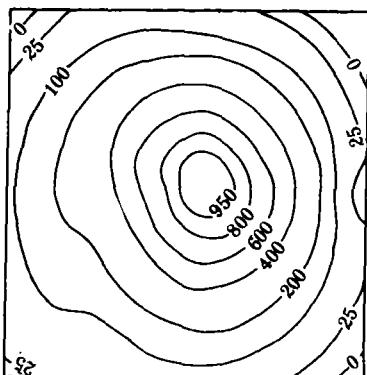


图 7 在图 6 中加入均方差为 1.659 nT 的随机噪声再经 KL 变换法去噪后的平面等值线图

方差为 1.659 nT 的随机噪声, 应用 KL 变换, 提取前三个特征 ($\sigma(3) = 1.6 < 1.659 \text{ nT}$), 重构数据, 从而得到平面等值线图 7。再与图 6 相比, 显见两者吻合程度显著, 从而说明了曲面上的重磁资料也可采用 KL 变换法进行去噪处理, 以便有利于“曲化平”工作。

3 实际资料处理

我们选用广西柳州地区 $1:20$ 万航磁测量结果(苏建明, 1981)经过频率域化极后的平面等值线图 8^①。由图可看出其异常形态比

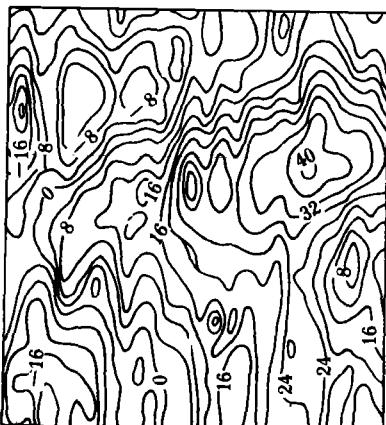


图 8 柳州地区 $1:20$ 万航磁测量结果经频率域化极后的平面等值线图(等值线间隔为 4 nT)

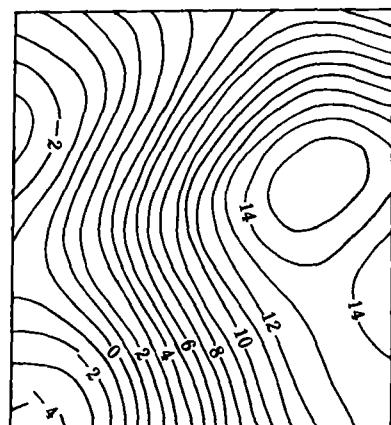


图 9 向上延拓 10 km 的平面等值线图
(等值线间隔为 1 nT)

较复杂, 高频干扰比较严重, 且呈南北向拉长形态, 其主要原因是化极场中误差的影响, 包括处理中夹杂着人为取数所带来的误差和测量结果的预处理中给位场值产生的一定误差等随机噪声。为了剔除这些噪声, 刘建军^①使用了一种低通滤波, 即向上延拓 10 km , 而得到图 9。从图 9 中可看出, 虽然其异常形态简单, 但其等值线图边部有畸变。这很可能是部分随机噪声传递所致。为此, 我们先采用 KL 变换方法去噪, 后向上延拓。

将图 8 以 50×40 的矩形网格划分, 即取

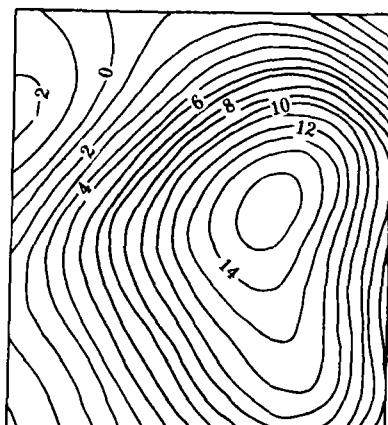


图 10 经 KL 变换后再向上延拓 10 km 的平面等值线图
(等值线间隔为 1 nT)

① 刘建军. 磁测数据处理一化极问题的研究. 西安地质学院硕士论文, 1992 年 6 月。

南北向剖面 40 条,每条剖面上取 50 个测点数据,构成一个 50×40 的数据矩阵。按 KL 变换求出数据协方差矩阵的特征值,由(4)式 $\eta(M)$ 曲线形态选取 $M=2$,即提取第一第二特征,再重构数据和向上延拓 10 km 即得到平面等值线图 10。对照图 9 可知,图 10 的等值线不仅形态简单,而且对称性较好,结果令人满意。

4 结语

a. 作者提出了 KL 变换方法在区域物探数据处理中剔除随机噪声的方法,其基本步骤是 KL 变换、提取特征和重构数据,建立了划分信号与噪声、提取信号特征的原则,使重构数据可以将随机噪声剔除到最低限度。

b. 通过人工资料和实际资料的处理,应用 KL 变换法能成功地从区域物探数据中剔除随机噪声,大大地增进了信噪比。

c. 在测区范围很大、剖面数和每条剖

面上测点数目很大($N, P \geq 50$)时,数据矩阵维数增高,考虑到一般的微机容量内存问题,我们可以采取分块处理办法进行 KL 变换法去噪,从而达到预期的效果。

在本文撰写过程中,地矿部物化探研究所的刘建军提供了有关实际资料,西安地质学院潘作枢教授提出过宝贵意见,在此深表谢意。

参考文献

- 1 K HSu. Wave separation and feature extraction of a-acoustic well-logging waveforms using Karhuen-Loeve transformation. *Geophysics*, 1990, 55, 176~184
- 2 郑治真. 数学信号处理基础. 北京: 地震出版社, 1988
- 3 程民德等著. 图象识别导论. 上海: 上海科技出版社, 1983
- 4 重磁资料数据处理问题. 北京: 地质出版社, 1977

(收稿日期 1992-12-02)

THE DIGITAL FILTERING OF KL TRANSFORMATION METHOD IN THE REGIONAL GEOPHYSICAL EXPLORATION

Xia Kewen (Logging Institute, Xi'an Petroleum Exploration Instrument Complex)

Abstract This paper presents the KL transformation method for the regional geophysical exploration data processing, and the method can reduce the ransom noise to the inferior limit by KL transformation, extract characterise and reconstruct data. The results of synthetic and practical data processing indicate that the digital filtering effect is very good and the signal-noise radio is fully increased.

Keywords numerical filtering; KL transformation method; geophysical exploration

第三届全国沉积学及岩相古地理 学术会议在成都召开

会议于 1992 年 10 月 5 日至 9 日召开。参加大会的有来自全国各地地质、石油、煤炭、冶金、有色金属、化工、核工业等行业的科研、生产、教学单位的 327 位代表,其中科学院学部委员 5 人。

会议收到论文 300 多篇,内容之广,数量之多,研究程度之深超过历届会议。会议学术交流反映出我国近几年沉积学和岩相古地理研究又有许多新进展,主要表现为:

a. 基础研究和应用研究继续深入,前缘研究有较大进展,而新技术、新方法研究尚需努力。

b. 岩相古地理研究,除进行基本的不同地层和

地区的岩相古地理分析外,还从不同角度,如生物群、遗迹化石、风暴作用、火山事件、古气候等方面进行系统综合分析。此外,还进行模拟和实验研究。

c. 应用研究方面的突出特点是结合找矿和国民经济建设需要,以找煤和找油气方面论文最多。在论述煤和油气形成及赋存机理时,既注意探讨和沉积相的关系,也注意构造和盆地演化的作用。

d. 层序地层和海平面变化研究是当前沉积学的热门课题。由于引进时间不长,尚在发展中。

(关世桥)