高等学校教学参考书

地球物理数据分析

——离散反演理论

王明光　楼　海　译

地质出版社

地球物理数据分析
——离散反演理论

William Menke 著
王明光 楼海 译
萧敬湧 王建谋 校

地质出版社

内 容 提 要

反演理论是近十几年发展起来的一个数学分支，它的出现大大推动了自然科学的诸多学科，包括地球物理学中数据分析和反演问题的求解。

本书是作者在哥伦比亚大学等校多年讲课的基础上编写而成的，既有理论又有实际应用，概念清晰，深浅适于大学高年级学生学习，是一本较好的入门教材。

全书共十二章，第1、2章提供一般的基础知识；第3～7章讲述高斯、线性反问题的解法和理论；第8、9章把讨论扩展到非高斯和非线性问题中；第10～12章进一步提供反演方法，提供应用实例及数值算法。附录中简单交待了复变量的反演理论。最后还附有英-中对照表。

译文文字通顺，并准确表达了原意。

本书适于应用地球物理专业作为“数据处理：离散反演理论”课程的教材，亦适于一切需要分析数据的自然科学工作者阅读参考。

Geophysical Data Analysis:
Discrete Inverse Theory

William Menke

Academic Press 1984

地球物理数据分析
——离散反演理论

王明光 楼 海 译
萧敬湧 王建谋 校

责任编辑 袁 方
地质出版社出版
（北京西四）
地质出版社印刷厂印刷
（北京海淀区学院路29号）
新华书店总店科技发行所发行

开本：850×1168 1/32 印张：8 字数：207,000
1988年4月北京第一版·1988年4月北京第一次印刷
印数：1—1,235 册 定价：1.80 元
ISBN 7-116-00144-1/P·128

前　言

每一位分析过数据的应用科学研究者都已经实践了反演理论。简言之，反演理论是用于从实测值中萃取关于研究对象有用推论的一组方法。用直线拟合一组数据就是反演理论的简单应用，而在医生们研究出CAT扫描设备后流行起来的层析成象技术则是反演理论在更高水平上的应用。

不过，对反演理论的研究不仅仅是把分析数据的方法罗列起来，而且还努力把这些方法条理化，找出它们潜在的相似性，撇开它们的差异；并研究从任何已知数据组中能够搜集到多少信息等基本问题。

物理特征可以大体上分为两类，即可以用离散参数描述的一类（如地球的质量或原子在蛋白质分子中的位置）和必须用连续函数描述的一类（如地球表面的温度或电容器内的电场强度）。对这两类不同的参数，在反演理论中使用不同的数学方法：对离散参数应用矩阵方程理论，而对连续函数则应用积分方程理论。

鉴于系入门性质的，本书仅涉及"离散反演理论"，即反演理论中与离散参数有关的那一部分。离散参数既可以是真正的离散值，也可以是适当离散化的近似值。只要不超出这些限制，本书讲叙的反演理论对于应用学科的大多数低年级研究生和很多高年级大学生来说是容易接受的。另外，本书仅假定读者具有微积分和线性代数的使用知识，并在一定程度上熟悉概率论和统计学中的一般概念。

本书对反演理论的论述分为四个部分。第1、2章提供一般的基础知识，即解释什么是反问题，其解是怎样构成的，以及复习本书将用到的概率论的某些基本概念。第3～7章讨论典型反问题的解，即具有高斯统计特征的线性问题的解。在所有反问题中对

这类问题认识得最彻底。同时，在这类问题中可以最清楚地引出不确定性、唯一性和分辨率等基本概念。第8、9章把讨论扩展到非高斯和非线性问题中。第10～12章提供反演理论的应用实例和数值算法的讨论。这些算法是在利用计算机求解反问题时必须用到的。

在本书的写作过程中作者得到过很多人的帮助。非常感谢作者在哥伦比亚大学和俄勒冈州立大学讲授反演理论课时参加听课的学生们,他们提出了很多有益的意见,Leroy Dorman，L. Neil Frazer和Walt Pilant仔细阅读了手稿，对他们所提的意见和所给予的鼓励表示感谢。对编辑Ellen Drake和制图员Susan Binder在他们工作中所做的努力也一并致谢。此外在写作本书时引用了很多位科学家和数学家的观点，在此也向他们表示感谢。

目　录

引　言

反演理论是一组条理化了的数学方法，用于根据从观测资料中得到的一些推断来简化数据，以便获得有关物质世界的有用信息。在本书中将要考虑的反演理论仅限于能够用数值表示的观测资料和问题。物质世界的观测资料是由测量值表或“数据”构成。我们要回答的问题将用代表物质世界特殊性质（但是未必可以直接测量）的数值（和统计特性）来陈述。由于这些特性将变得明显，因而称为“模型参数”。我们将假设存在把模型参数和数据联系起来的某些特殊方法（通常是一个数学公式或数学模型）。

例如，“什么使行星运动”就不是应用反演理论所能回答的问题。尽管它确实具有科学和历史的重要性，但是它的答案从性质上讲却不是数字形式的。相反，假设应用牛顿力学，根据观测到的哈雷彗星轨道来确定行星的个数及轨道却是可以应用反演理论的问题。行星的个数及其轨道运行位置表从性质上讲是数字形式的。这两类问题的另一重要差别是，前者是要确定沿轨道运动的原因；而后者是预先假定造成沿轨道运动的原因，然后仅仅是确定某些细节。反演理论难以提供第一类问题所要求的那类解答；因为要回答这类问题往往要求事先确定物理模型。

“反演理论”这一术语是与“正演理论”相对应的。**正演理论**的定义是，根据某些一般原理或模型，以及一系列与所处理的问题有关的已知具体条件来预测观测结果（预测数据）的方法。而**反演理论**，粗略地讲，就是处理相反的问题，即从数据及某些一般原理或模型出发来确定模型参数的估计值。在上面的例子中，根据推测的天体运行位置表来预测哈雷彗星的轨道，就是一个有关正演理论的问题。

再以地下温度变化是深度的函数这一现象，来比较正问题和反问题。假设地球内的温度随深度呈线性增加，也就是温度T与深度 z 的关系式为$T=az+b$，其中的a和b是数值常数。如果已知$a=0.1$和$b=25$，那么对于任何深度都能利用公式直接计算温度，即解正问题。反问题则是根据在钻孔中不同深度实测的一组温度值来确定参数a和b。可以看到，这是一个用直线拟合一组数据的问题。它比计算一个一阶多项式的值这样的正问题难得多。这是大多数反问题的共同特点，即解反问题比解相应的正问题明显地难得多。

正问题： 模型参数→模型→数据的预测

反问题： 数据→模型→模型参数的估计值

注意，反演理论的主要目的是提供有关模型中的未知数值参数的信息，而不是提供模型本身。尽管如此，反演理论还是经常能够提供一种方法来评价给定的模型正确与否，或者在几个可能的模型中辨别何者正确。

反演理论中所遇到的模型参数既可以是离散值，也可以是一个或多个变量的连续函数。上述直线的截距和斜率就是离散参数的例子；而随位置不同温度发生连续变化，则是连续函数的例子。本书仅涉及模型参数可以用一组有限个数值来表示的离散反演理论。实际上，这一限制并没有排除对连续函数的研究，因为连续函数通常能够用有限个离散参数适当地逼近。例如，温度可用有限个彼此相隔很近的点上的值来表示，或者用一组具有有限个系数的样条表示。不过，这一近似达不到利用连续函数进行研究的精度。因为连续函数的参数化总是近似的，而且在某种程度上具有随意性，这就给反演理论带来了一定的不精确性。尽管如此，离散反演理论通常仍然是研究反演理论的良好出发点，这是因为它主要依靠向量和矩阵理论，而不是依靠某种更加复杂的连续函数和算子理论。不仅如此，认真地使用离散反演理论经常能得到很有价值的见解，即使应用到某些包含连续参数的问题上也是如此。

尽管反演理论的主要目的是提供模型参数的估计值，但是它却有着相当大的适用范围。甚至在模型参数是唯一拟求的结果之情况下，也存在大量可供提取的有关信息，可用来帮助确定反问题解的“优度”。而当我们主要是在实验设计或数据的概括中把反演理论作为一种工具来使用时，模型参数的真实值就显得不太重要了。下面是一些反演理论可以帮助回答的问题：

（a）反问题之间的根本相似性是什么？

（b）如何得到模型参数的估计值？

（c）模型参数估计值中的误差有多少来自测量值的误差？

（d）给定一个特定的实验设计就能真正确定某一组模型参数吗？

这些问题强调的是，反问题存在许多不同种类的答案和许多用来评价这些答案“优度”的准则。反演理论的许多研究课题都涉及到识别什么时候某些准则比其它的准则更适用，以及发现并进而避免（如果有可能的话）可能出现的各种各样的失误。

反问题存在于自然科学的许多分支之中，据不完全统计有以下几个方面：

（a）医学上的层析成象，

（b）图象增强，

（c）曲线拟合，

（d）震源定位，

（e）因子分析，

（f）根据地球物理数据确定地球结构，

（g）卫星导航，

（h）利用干涉法测绘宇宙射电源，

（i）利用 x 射线的衍射分析分子结构。

反演理论是由具有各种各样背景和目的的科学家和数学家们发展起来的。因此，尽管对这一理论的各种叙述（version）具有坚实的和根本的相似性，但是表面上看来却差别很大。本书的目标之一，是以个别观点和“大概念”两者都能被清楚理解的方

式来介绍离散反演理论的各个方面。

从对反演理论的探讨来看，大概有三种主要观点。第一种，也是最早的一种观点，来自概率论——把"噪音"量作为观察现实世界的一个自然出发点。在反演理论的这一提法中，把数据和模型参数作为随机变量来处理，而且强调很多的是确定它们所服从的概率分布。这一观点很自然地导致对误差的分析和对答案意义的检验。

第二种观点是从自然科学中保持确定状态和避免直接应用概率论的部分发展起来的。这种方法倾向于仅涉及模型参数的估计值（而且多半带有其误差范围）而不涉及概率分布本身。然而，一个估计值常常正是概率分布的期望值；其差别只不过是强调哪一个而已。

第三种观点产生于模型参数本身是连续函数的考虑。前两种观点对这一问题的处理，是把连续函数用有限个离散参数来近似，而已经发展起来的第三种方法则是直接对连续函数进行处理。尽管连续反演理论不是本书讨论的内容，但是许多最初由连续反演理论发展起来的概念已经在离散反演理论中得到了应用，特别是应用于离散化了的连续函数。

本书可以作为一年级研究生学习反演理论的教材。虽然反演理论是一门数学课程，但是本书尽量保持完整的数学处理。除了个别情况外，一般假定读者仅具备微积分和矩阵代数知识。然而，这样处理并不意味着简单化。从其它科学文献上找出了一些实例来说明各种不同的方法。由于在实践中大多数反问题的求解都需要进行大量的计算，因而本书对于可以在现代计算机上实现的各种算法都给予了一定的注意。

第一章　对反问题的描述

1.1　反问题公式的建立

在大多数反问题中，一开始都是对数据进行描述。因为多数情况数据就是一个简单的数值表，所以这些数值可以很方便地用向量来表示。例如，如果在某个特定的实验中得到了N个测量值，就可以考虑把这些数值表示成维数为N的一个向量 $\boldsymbol{d}$。类似地，模型参数可用维数为M的向量$\boldsymbol{m}$表示，即

$$\text{数据：}\ \boldsymbol{d}=[d_1,d_2,d_3,d_4,\cdots,d_N]^{\mathrm{T}}$$

$$\text{模型参数：}\ \boldsymbol{m}=[m_1,m_2,m_3,m_4,\cdots,m_M]^{\mathrm{T}} \tag{1.1}$$

其中的上角标T表示转置。

对一个反问题的基本陈述就是把模型参数和数据以某种方式联系起来。这一关系被称之为模型，它通常为一个或几个公式，数据和模型参数满足这些公式。

例如，若通过测量物体的体积和质量来确定其密度，则有两个数据——质量和体积（比方说分别为d_1和d_2）和一个未知的模型参数——密度（比方说m_1）。该模型可以叙述为密度乘体积等于质量，并且可以严密地表示成向量方程$d_2m_1=d_1$。

在较多的实际情形中，数据和模型参数以更复杂的方式联系着。最一般的情况是，数据和模型参数由以下的一个或多个隐式方程联系起来，如

$$\begin{aligned} f_1(\boldsymbol{d},\boldsymbol{m})&=0\\ f_2(\boldsymbol{d},\boldsymbol{m})&=0\\ &\vdots\\ f_L(\boldsymbol{d},\boldsymbol{m})&=0 \end{aligned} \tag{1.2}$$

式中的L是方程的个数。在上述有关密度测量的例子中，$L=1$和$d_2m_1=d_1$可以构成一个形如$f_1(\boldsymbol{d},\boldsymbol{m})=0$的方程。这些隐式方程可以严密地表达成向量方程$\boldsymbol{f}(\boldsymbol{d},\boldsymbol{m})=0$，它概括了对测量数据和未知模型参数有何关系的了解程度。反演理论的目的就是求解或“变换”这些模型参数方程，或者回答在任何给定的情形哪种类型的答案是可能的或合乎需要的。

对于方程$\boldsymbol{f}(\boldsymbol{d},\boldsymbol{m})=0$，究竟是包含了足够多的可用以唯一确定模型参数的信息，还是其本身就是相容的，并未做任何声明。反演理论的目的之一就是解这类方程并提供处理其中的问题的方法。通常，$\boldsymbol{f}(\boldsymbol{d},\boldsymbol{m})=0$可能由数据和模型参数的任意复杂的（非线性）函数组成。然而，在许多问题中，这一方程呈某一简单形式（几种简单形式之一）。因为这些特殊情形在实际问题中经常出现，而且在后面的章节中还要对它们进行专门讨论，所以为方便起见，对其中一些给予命名，如下所述。

1.1.1 隐线性形式

函数$\boldsymbol{f}$对数据和模型参数来说都是线性的，因而可以写成矩阵方程的形式，如

$$\boldsymbol{f}(\boldsymbol{d},\boldsymbol{m})=0=\mathbf{F}\begin{bmatrix}\boldsymbol{d}\\ \boldsymbol{m}\end{bmatrix} \tag{1.3}$$

式中的$\mathbf{F}$是一个$L\times(M+N)$阶矩阵。

1.1.2 显形式

在许多情况下，有可能把数据和模型参数分开，并形成$L=N$个对数据是线性的方程（但是向量函数$\boldsymbol{g}$仍然是模型参数的非线性函数），即

$$\boldsymbol{f}(\boldsymbol{d},\boldsymbol{m})=0=\boldsymbol{d}-\boldsymbol{g}(\boldsymbol{m}) \tag{1.4}$$

1.1.3 显线性形式

对于显线性形式，函数$\boldsymbol{g}$也是线性的，这样就得到了一个$N\times M$阶矩阵方程（其中$L=N$）

$$\boldsymbol{f}(\boldsymbol{d},\boldsymbol{m})=0=\boldsymbol{d}-\mathbf{G}\boldsymbol{m} \tag{1.5}$$

使用这一形式，等于说1.1.1节中的矩阵$\mathbf{F}$是一个分块对角

阵

$$\mathbf{F}=\begin{bmatrix} -\mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{G} \end{bmatrix} \tag{1.6}$$

1.2 线性反问题

最简单和最易理解的反问题是那些能用显线性方程 $\mathbf{G}\boldsymbol{m}=\boldsymbol{d}$ 表示的问题。因而该方程成了研究离散反演理论的基础。正如在下面将要看到的，自然科学中出现的许多重要反问题确实包含着该方程。而其它的一些，当包含着更复杂的方程时，则常常能够通过线性近似来求解。

类似于积分方程理论，称矩阵$\mathbf{G}$为*数据核*，只是在积分方程理论中数据和模型参数向量是两个连续函数$d(x)$ 和 $m(x)$，其中的 x 是某一独立变量。这两个函数由下面的方程联系起来

$$d(x)=\int G(x,\xi)m(\xi)d\xi \tag{1.7}$$

式中函数 $G(x,\xi)$ 是积分方程的核，或者称为格林函数。这类问题的求解属于连续反演理论的范围。

1.3 建立反问题公式的例子

1.3.1 例1：直线拟合

假定在地下 z_i 处得到N个温度测量值T_i。这样，数据就是由N个温度测量值组成的一个向量$\boldsymbol{d}$，即$\boldsymbol{d}=[T_1,T_2,T_3,\cdots,T_N]^{\mathrm{T}}$。严格地说，深度 z_i并不是数据。但是它们却可以提供一些辅助信息来描述实验的几何特征。这一差别将在后面得到进一步的阐明。

假定在某一模型中，温度是深度的线性函数，即$T=a+bz$。那么截距 a 和斜率 b 就成了该问题的两个模型参数，$\boldsymbol{m}=[a,b]^{\mathrm{T}}$。根据模型，每个温度观测值必须满足$T=a+bz$，也就是

$$
\begin{aligned}
T_1 &= a + bz_1 \\
T_2 &= a + bz_2 \\
&\vdots \\
T_N &= a + bz_N
\end{aligned}
\tag{1.8}
$$

这些方程可以整理成矩阵方程$\boldsymbol{Gm}=\boldsymbol{d}$形式

$$
\begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_N \end{bmatrix} = \begin{bmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_N \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}
\tag{1.9}
$$

1.3.2 例2：抛物线拟合

如果把例 1 中的模型改变一下，假定温度随深度呈形如$T=a+bz+cz^2$ 的多项式关系，那么就要对该问题加上一个新的模型参数，$\boldsymbol{m}=[a,b,c]^{\mathrm{T}}$。现在模型参数的个数为 $M=3$。假如数据满足

$$
\begin{aligned}
T_1 &= a + bz_1 + cz_1^2 \\
T_2 &= a + bz_2 + cz_2^2 \\
&\vdots \\
T_N &= a + bz_N + cz_N^2
\end{aligned}
\tag{1.10}
$$

这些方程可以整理成矩阵方程

$$
\begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_N \end{bmatrix} = \begin{bmatrix} 1 & z_1 & z_1^2 \\ 1 & z_2 & z_2^2 \\ \vdots & \vdots & \vdots \\ 1 & z_N & z_N^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}
\tag{1.11}
$$

这一矩阵方程具有形如$\boldsymbol{Gm}=\boldsymbol{d}$的显线性形式。但是应该注意到，尽管这一方程对数据和模型参数来说都是线性的，但是对辅助变量 z 来说，却不是线性的。

该方程在形式上非常类似于例 1 中的方程，这就是采用矩阵表示法的根本原因之一：因为它经常能够强调表面上不相同的问题之间的相似性。

1.3.3 例3：声学层析成象

假定把一些砖按正方形排列成一堵墙（图1.1），其中的每一块砖由不同类型的粘土做成。如果不同的粘土其声速不同，通过