反 演 理 论

数 据 拟 合 及 模 型 参 数 估 算 方 法

艾伯特·塔兰托拉 著

张先康等 译

刘福田 校

学术书刊出版社

内容提要

本书较系统地介绍了反演问题的理论，论述了离散反演问题和一般反演问题，以丰富的实例说明了反演理论在地球物理学科中的应用，内容涉及新近发展起来的地球物理层析成象方法。本书对从事地球物理学研究以及其它有关学术领域研究人员有很好的参考价值。

• • •

INVERSE PROBLEM THEORY

Methods for data fitting and model parameter estimation

Albert Tarantola

Elsevier Science Publishers B·V·, 1987

• • •

反 演 理 论

数据拟合及模型参数估算方法

艾伯特·塔兰托拉 著

张先康等 译 刘福田 校

责任编辑：姜维岐

•

学术书刊出版社出版（北京海淀区学院南路86号）

新华书店北京发行所发行 各地新华书店经售

北京印刷三厂联营厂印刷

•

开本：850×1168毫米 1/32 印张：$17\frac{3}{8}$ 字数：453千字

1989年12月第1版 1989年12月第1次印刷

印数：1—1500册 定价：9.50元

ISBN 7-80045-509-2/P·14

目　录

第一部分　离散反演问题

第二部分 一般反演问题

第一章　一般离散反演问题

针对正确问题的近似答案，即使往往是模糊的，也要比针对错误问题所能精确得到的确切答案要好得多。

J.W.Tukey，1962.

本章的中心内容是一组参数集的"信息状态"这一概念。假定，描述一参数集信息状态最普遍的方法是在相应参数空间中定义一概率密度函数，于是，可观测参数的测量结果（数据），模型参数的先验信息，以及可观测参数与模型参数间物理上相关联的信息均可用概率密度来描述。因而，一般反演问题可视为所有这些信息"组合"的问题。由这里提出的观点，反演问题的求解以及误差和分辨率的分析，可由完全非线性方法来完成（但也许要化长得惊人的计算时间）。一般场合下，该方法的结果可化简为更常规方法的结果。后面各章的所有结果均可由本章的论证证明为合理的。

1.1　模型空间和数据空间

设所研究的物理系统为S。譬如，S对天文学家来说可能为一星系，对地球物理学家来说可能为地球，对量子物理学家来说可能为一量子。

研究一物理系统的科学过程可划分为（相当人为的）如下3个步骤：

i）系统的参数比：寻求最小数目的模型参数，其数值可完全描述系统特征（从一给定观点出发）。

ii）正演模拟：寻求物理定律，由所给的模型参数值预测可测参数的测量结果。

iii）反演模拟：用可测参数的某些实测结果推断模型参数的真值。

这些步骤间存在着强反馈作用，其中某一步骤的惊人进展往往会导致其余两步骤的进展。

前两步骤主要为归纳法，第三步骤则主要为演绎法。这意味着在前两步骤我们所沿用的假定和思想方法难以用显式表达。相反，逻辑数学理论（由概率论实现）似乎相当好地适用于第三步骤。该步骤为本书所致力探讨的问题。

1.1.1 模型空间

一般来说用以描述一个系统的模型参数的选择是不唯一的。

例1 为描述固体的弹性性质，可用每一点X上将应力$\sigma^{ij}(x)$和应变$\varepsilon^{ij}(x)$联系起来的弹性刚度张量$c^{ijkl}(x)$

$$\sigma^{ij}(x)=c^{ijkl}(x)\varepsilon^{kl}(x)$$

另一方面，也可用将应变和应力联系起来的弹性柔度张量$s^{ijkl}(x)$

$$\varepsilon^{ij}(x)=s^{ijkl}(x)\sigma^{kl}(x)$$

其中张量s为c的逆：

$$c^{ijkl}s^{klmn}=\delta^{im}\delta^{jn}$$

运用刚度或是运用柔度完全等价，并不存在“固有的”选择。

模型参数特定的选择称为系统的参数化。如果两种不同的参数化是以双射相联系，那么它们是等价的。

可以不依赖于任何特定的参数化方法而引进一个点的抽象空间集合，每个点表示系统的一想象的模型。该空间称为模型空间并以$\mathcal{M}$来表示。

为了对系统进行定量地讨论，须选用一特定的参数化方法。定义一参数化方法意味着定义一组实验过程，从而能够测定（至少在理论上）系统的不同特征。一特定的参数化方法一旦选定，模型空间中的每一点将与一组数值相联系，用M空间的一点来表示，M同构于R^n的一个部分（R表示实线，n为参数的个数）。模型空间$\mathcal{M}$被本征地定义，而M空间则取决于所选择的具体参数化方法。

从数学观点来看，$\mathcal{M}$为一（非线性）流形，而M则为$\mathcal{M}$的一

个图。

例2：当一核爆炸发生于地球表面时，它产生在地球内部传播的地震波。因而可用一些地震台站的波到时估计爆炸的位置。于是，一个“模型”为爆炸的一个特定位置，亦即地表的一个几何点。模型空间$\aleph$可用一单位球的球面来表示（图1.1左）。模型空间就被本征地定义了，即并没有参照任何特定的坐标系。然而，为了进行数值计算，有必要将模型空间的每个点用给定的坐标系中坐标来表示。对模型空间定义一坐标系就是给出模型空间的一个图M。于是每个模型相应于R^2中的一部分的一个点（图1.1右）。本例中“图”这个词还可视为原来字面上的意义。经引伸后，数学家们将n维非线性流形和R^n的一部分间的任何“映射”均称作“图”。模型空间的一个映象也经常被叫做“模型空间”，这是语言上的误用。这种误用唯一的不便在于，它将导致这样的印象，似乎模型空间必须是线性空间，但一般来说并非如此。在此特例中，可为模型空间引入一个任意两点间的距离的概念（它既可为连接球面两点穿过球的直线段的欧几里德长度，亦

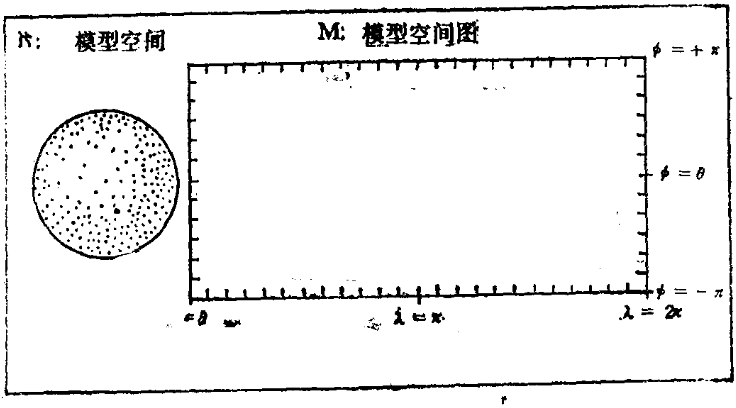

图1.1 模型空间$\aleph$一般为一非线性流形（左图）。为进行数值计算，我们引入流形的图（右图）。用不严格的词语表达，仍称其为“模型空间”（见正文的讨论）

可为连接球面两点大圆弧的长度）。这属一特例而并非普通规律。对于一般的模型空间，不可能以这么自然的方式引入距离的概念。

对微分流形理论感兴趣的读者可参考Lang（1962）、Naasimhan（1968）或Boothby（1975）等人的著作。

我们已知道模型空间$\mathcal{M}$的图M与R^n“同构”。n维的图与R^n的差别在于，R^n的每个点为一组纯实数，即无量纲的量。相反，图M的每个点为给定物理量纲（压强、温度、电量……）的n个值。

$\mathcal{M}$的每个点、或M的相应的点，被称为一个模型，并以m表示。

给定一模型空间$\mathcal{M}$。如果能不致于引起混乱，则用非严格的说法，该空间的一个特定的“图”仍被称之为“模型空间”。这相应于反演问题文献中的传统叫法。

完全地描述一个系统所需模型参数的数目既可是有限的，亦可是无限的。

例3 例1中若固体的弹性性质逐点变化，那么，为了完全描述该系统，我们就需要无限个值。若固体为均匀的（即弹性参数的值不依赖于x），那么，21个参数就足以对它全面描述。

令m^α代表（一组有限或无限的参数中）一特定的参数。参数m^α的取值可为离散的或连续的。例如，若m^α表示太阳的质量，则可先验地假定它可取零至无穷的任何值；若m^α表示量子的自旋，可先验地假定它仅能取离散的值。以后我们将会看到，通过“δ函数”的运用，就可以将离散参数视为连续参数的特殊情况。为简化讨论，本书所用的术语均相应于假定所考虑的参数取其连续值。如在个别问题中不是这种情况的话，读者很容易作出相应的修改（见问题1.3和1.4）。

无限维空间理论比有限维空间理论需要更多的技术词汇。在下文，即本书的第一部分，假定模型空间是有限维的。按照数学的观点，将系统限制为有限个参数是相当苛刻的，因为函数领域（即使是连续的）比有限维空间领域广泛得多。但是，只要在一些问题中我们能够承认所考虑的函数是有限带宽的，则总可考

虑对这些函数的采样。在这种情况下，对反演问题理论来说，函数方法和离散方法所得的结果没有差别（但是，数值方法可能差别很大，如对比问题1.2和7.1即可发现）。

当系统的特定的参数化方法选定之后，每个模型可用一组特定的模型参数值来表示：

$$m=\{m^{\alpha}\} \qquad (\alpha\in I_M) \qquad (1.1)$$

其中I_M为一组有限个离散角标值。只要S的一个特定的参数化方法可视为对$\mathcal{M}$选择坐标线，变量m^{α}即可称为m（相对于给定出坐标线）的坐标值。

根据我们新定义的术语，下列提法为等价的：

i）对物理系统S参数化，

ii）对模型空间$\mathcal{M}$定义一坐标系，

iii）定义$\mathcal{M}$的一个图M

根据定义，$\mathcal{M}$的图与R^n同构。一个特例是图为线性（向量）空间。那么，对给定的图M就可定义两模型m_1与m_2之和为其“分量”之和

$$(m_1+m_2)^{\alpha}=m_1^{\alpha}+m_2^{\alpha} \qquad (\alpha\in I_M) \qquad (1.2)$$

模型与一实数相乘为与其“分量”相乘，

$$(rm)^{\alpha}=rm^{\alpha} \quad (\alpha\in I_M) \quad (r\in R) \qquad (1.3)$$

这表明称坐标m^{α}为“分量”是合乎道理的。然而，应该强调，上述定义并非是本征的，因为这样定义的和或积依赖于S的参数化方法。一般来说给出m_1+m_2或rm的本征定义是不可能的，因模型空间$\mathcal{M}$一般情况下并非线性空间。

例4 例1中的均匀固体可用21个弹性刚度系数c^{ijkl}或21个弹性柔度系数s^{ijkl}来描述。设m_1和m_2表示两不同的弹性固体，他们可用c_1^{ijkl}和c_2^{ijkl}或者s_1^{ijkl}和s_2^{ijkl}来表示。由$c_1^{ijkl}+c_2^{ijkl}$定义的弹性体与$s_1^{ijkl}+s_2^{ijkl}$所定义的弹性体不同。两模型的和不能被本征地定义。

事实上，反演理论完全可不用参照任一特定的参数化方法而发展起来。下面我们将会看到，为了处理反演问题最普遍的公式，

唯一要确定的数学目标是模型空间的测度（积分意义下的）。$\mathfrak{M}$的测度为一映射。它把$\mathfrak{M}$任一子集A与一正实数相联系。

$$A \rightarrow P(A) \in R^{+}$$

称为A的测度。一般$P(\mathfrak{M})=1$，且此测度称为$\mathfrak{M}$上的概率。从理论上讲，该测度可与$\mathfrak{M}$的任何参数化方法无关地定义，即是说，不依赖于任何特定的图。但是，一特定的图M一旦选定，那么，就非常容易用概率密度法描述概率。

实际应用中，我们遇到的总是一特定的参数化方法。该方法一经选定，我们既可忽略抽象的模型空间与其图的细微差异，而称后者为模型空间，即可忽略线性向量结构为非本征的，并进行计算。

对一物理系统定义参数化时，曾谈到须定义一些参数以便于“至少在理论上”可以度量。下面的例子说明了这一点。

例5 地球金属核的半径可“实验性地”定义为地球内这样一个球体的半径，在该球体内矿物成分大部分是金属。我们知道如何确定给定样品的化学成分。钻一（相当）深的孔并分析得到的样品，就“足以”测量地球核的半径。“核半径”这一参数就这样很完善地被定义了下来，尽管仅仅在理论上为可测的。运用物理定律我们可预测到达核表面的地震波的特性。在地表可直接观测到衍射——反射波。于是，反演理论可使我们得到核半径的信息。

1.1.2 数据空间

为得到模型参数的信息，我们须在一物理实验过程中进行一些观测。也就是说，我们须对一些可测参数进行测量。

例6 如例5中所讨论的，对一个兴趣在于地球物理深部构造的地球物理学家来说，观测量可能为在地表记录的一组地震图。

例7 对一个粒子物理学家来说，观测量可能为入射粒子流在靶上沿不同角度散射的通量。

实验者的任务是困难的，这不仅由于他必须尽可能地测量得

准确，更重要的是，他必须设想新的实验方法使得他所观测的量携带模型参数的最大限度的信息。例如：确定船长的年龄可能很容易（例如，可通过偷窃他的护照），但是很少有这样的机会，这种测量会携带船上桅杆数目的许多信息。

与模型参数的情况一样，实验设备给定后选择可测参数时仍有一定的自由度。

例8 对于一地震仪，我们可以选择正比于地动位移、速度或加速度的电压信号作为它的“输出”。

于是，我们得到了数据空间的抽象概念。它可定义为所有可能的仪器响应所组成的空间。每个特定的实现用d 来表示。当“可测参数”一经选定后（如例8），数据空间的一个图，以D 表示。用不太严格的语言描述，仍称其为“数据空间”。任一可能的测量结果可写成“分量”。

$$d=\{d^i\} \qquad (i\in I_D) \tag{1.4}$$

其中I_D表示一组离散的（且有限的）角标。通过定义

$$(d_1+d_2)^i=d_1{}^i+d_2{}^i \qquad (i\in I_D) \tag{1.5}$$

$$(rd)^i=rd^i \qquad (i\in I_D)\ (r\in R) \tag{1.6}$$

则D为线性空间。于是，每个向量d被称为数据向量，或数据组。

1.1.3 联合空间D×M

有时引入积空间X＝D×M是很有用的。其元素为一对x＝(d, m)。由于d的元素叫做观测参数，m的元素叫做模型参数，因而x的元素可叫做物理参数，或简称为参数。X空间于是被称为参数空间。

X可被直观地理解为还包括测量仪器本身的广义物理系统S。这个空间比D和M要基本得多。事实上，许多问题中将X分离为数据空间和模型空间可能是随意的。

x的分量可视为参数空间的坐标。我们可通过如下双射任意地定义另一坐标系：

$$x^*=x^*(x) \qquad x=x(x^*) \tag{1.7}$$

这样的坐标系称为等价的。在1.2.3节我们将假定坐标系是最小的。

m的分量用希腊字母α，β，……来表示，d的分量用小写拉丁字母i，j……表示。x的分量在必要时用大写拉丁字母表示：

$$x = \{ x^A \} \qquad (A \in I_X) \qquad (1.8)$$

1.1.4 符号

处理离散反演问题的作者往往用到列矩阵记法表示模型空间或数据空间的元素：

$$m = \begin{pmatrix} m^1 \\ m^2 \\ \cdots \end{pmatrix}$$

$$d = \begin{pmatrix} d^1 \\ d^2 \\ \cdots \end{pmatrix}$$

本书中，“向量”一词总是意味“线性向量空间的元素”，从不认为向量自然表示为一列矩阵。

例9 如离散的模型相应于二元函数f（x，y）的离散化，很自然地，模型空间的元素可表示为：

$$m = \begin{pmatrix} m^{11} & m^{12} & \cdots \\ m^{21} & m^{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} = \begin{pmatrix} f(x^1,y^1) & f(x^1,y^2) & \cdots \\ f(x^2,y^1) & f(x^2,y^2) & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}$$

因而，角标组I_M为N×N（即一组整数对）。这种场合下，二维（数组即矩阵）可方便地表示模型，而将m的分量排成一列矩阵却很不直观（在数值计算中向来也不需要）。

如果向量的分量须显式表达，则采用如下抽象的记法：

$$m = \{ m^\alpha \} \qquad (\alpha \in I_M)$$

$$d = \{ d^i \} \qquad (i \in I_D)$$

以上两式中并未假定任何特定的排列方式。

例如，令G表示从M到D的线性算子。我们写作：

$$d = Gm \tag{1.9a}$$

既然G将两离散空间联系起来，可以看出存在着常数$G^{i\alpha}$（$i \in I_D$）（$\alpha \in I_M$）使得线性方程（1.9a）可写为：

$$d^i = \sum_{\alpha \in I_M} G^{i\alpha} m^\alpha \qquad (i \in I_D) \tag{1.9b}$$

像G 那样的线性算子的分量一般可用多维数组（不要求 一 定为“矩阵”）表示。

例10 对于上例中的模型空间，有

$$d^i = G^{i11} m^{11} + G^{i12} m^{12} + \cdots$$

其中，角标i可依次为多维的。

当然，只要m和d 的分量个数为 有限的，我们即可将其重排为列矩阵，像G那样的线性算子可用 一般的二维矩阵来表示。但是，这样作的唯一效果一般是破坏了问题的所有对称性，并且引进了矩阵的运算。在这些问题中，稍微抽象一些的代数即可简化记号和运算（例如，见问题1.2）。作为一个固有的术语，表示线性算子的常数组应叫做算子的核，不能与算子本身等同起来。算子是更基本的概念（对一给定的线性算子，不同的核有如不同的基那么多，而基可在相应的线性空间中选择）。

令ϕ^α（或ψ^i）为常数使得$\Sigma m^\alpha \phi^\alpha$（或$\Sigma d^i \psi^i$）有意 义（尤 其是物理量纲的一致性），并给定一（无量纲的）实数。定义

$$m^t \ = \sum_{\alpha \in I_M} m^\alpha \phi^\alpha \tag{1.10a}$$

$$d^t \psi = \sum_{i \in I_D} d^i \psi^i \tag{1.10b}$$

在第4章和第5章，ϕ（或ψ）被视为M（或D）的对偶空 间 的 元素。

给定方程（1.9）的线性算子G，将模型空间映射到数 据 空间，用G^t表示G的转置，它为由下面恒等式定义的线性算子

$$m^t (G^t \psi) = (Gm)^t \psi \tag{1.11a}$$