现代数学译丛 33

统计与计算反问题

〔芬〕Jari Kaipio　Erkki Somersalo　著
刘逸侃　徐定华　程　晋　译

科学出版社

现代数学译丛 33

统计与计算反问题

〔芬〕 Jari Kaipio　Erkki Somersalo　著

刘逸侃　徐定华　程　晋　译

科　学　出　版　社

北　京

图字：01-2018-4306 号

内 容 简 介

本书主要介绍了处理反问题 (不适定问题) 的统计方法, 尤其侧重于建模与计算这两大问题. 与经典文献中处理反问题的方法不同, 本书立足于 Bayes 统计学的框架, 将所有变量都视作随机变量, 并把反问题的解以概率密度函数的形式给出. 同时, 对于数学模型本身存在的误差和数值离散导致的额外误差, 本书还创造性地进行了源自建模误差的统计分析.

本书详细讨论了先验模型的构造、测量噪声建模、Bayes 估值以及非静态统计反演方法等, 并引入 Markov 链 Monte Carlo 方法以及最优化方法来探究概率分布. 另外从 Bayes 统计学的角度重新研究了经典正则化方法, 揭示了两者之间的关系. 对于书中得到的结论和涉及的技法, 作者还佐以易懂但深刻的例子帮助读者理解. 本书将统计方法应用到一些较为前沿的问题中, 例如离散误差分析、模型降阶等. 在书中, 这些统计方法还被进一步应用于一系列实际问题中, 包括有限角度断层成像、图像去模糊、电阻抗断层成像、生物磁学反问题等.

本书主要面向应用数学、计算数学、统计学专业的教师、研究生和高年级本科生. 同时, 也适合具有扎实数学功底的工科、物理学和经济学专业的教师和研究生阅读. 此外, 亦可供各相关行业工作者在生产实践中参考.

图书在版编目(CIP)数据

统计与计算反问题/(芬) 亚里 · 凯拉 (Jari Kaipio), (芬) 埃尔基 · 索梅尔萨洛 (Erkki Somersalo) 著; 刘逸侃, 徐定华, 程晋译. —北京: 科学出版社, 2018. 8
书名原文: Statistical and Computational Inverse Problems
ISBN 978-7-03-058181-5

Ⅰ. ① 统… Ⅱ. ① 亚… ② 埃… ③ 刘… ④ 徐… ⑤程… Ⅲ. ① 统计方法 Ⅳ. ① C81

中国版本图书馆 CIP 数据核字 (2018) 第 139469 号

责任编辑: 李 欣 / 责任校对: 彭珍珍
责任印制: 张 伟 / 封面设计: 耕者设计

科 学 出 版 社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

北京教图印刷有限公司 印刷
科学出版社发行 各地新华书店经销
*
2018 年 8 月第 一 版 开本: B5 (720 × 1000)
2018 年 8 月第一次印刷 印张: 19 1/2
字数: 393 000

定价: 138.00 元
(如有印装质量问题，我社负责调换)

中文版序言

当前, 我们步入了一个数学深度发展和应用融合的新时代. 这个时代的科技发展呈现大数据深度应用、人工智能快速发展的崭新特征. 在这个新时代, 数学与统计学、数学与大数据、数学与人工智能等诸多领域深度融合, 孕育新发现、新理论、新方法、新技术!

在数学及相关学科, 反问题的研究快速发展, 成为了应用数学与计算数学的重要领域, 也成为工程技术中识别、控制、设计问题得以解决的核心技术!

反问题研究既可来源于具有确定性数学模型, 也可大量来源于具有随机规律的数学模型. 统计学正突破传统的应用范围向各学科渗透, 特别是反问题研究及应用领域.

本书英文原版成书于 10 年前经典反问题方法发展成熟之时, 最先将 Bayes 统计学的思想引入反问题的研究, 具有超凡的前瞻性. 此后, 越来越多的反问题学者将目光投向统计和随机方法, 本书的基本内容和精神实质依然具有重要的指导意义. 本书起到了 "连接经典和统计反演方法之桥梁" 的作用, 并提供了从统计学观点加深理解传统思想的视角.

相较于国外同行, 我国在统计反问题的研究上起步较晚, 从事该研究的学者还不多, 成果的数量和质量都与国际前沿存在不小的差距. 因此, 翻译、引进本书能填补该方向在国内反问题学界的空白, 对于推动反问题研究及其应用、应用数学的发展都将发挥巨大的作用, 均具有重要的学术价值.

本书以数学建模和数值反演为两翼, 将一系列统计方法应用于一些具有很强现实意义的实际问题中, 其中包括参数识别、图像处理、医学成像等相关学科中的重大课题. 可以预见, 如能将本书中的方法运用到这些领域的生产实践中, 则将对国民经济发展起到关键的作用.

国内外以经典视角研究反问题的书籍浩如烟海. 从计算的角度来看, 这类书籍通常假设测量误差很小, 且描述现象的模型精确已知, 而反问题的解也是确定值. 本书着眼于实际问题, 将观测数据、未知量乃至模型本身都设为随机变量, 并用随机性反映对其数值的可信度; 反问题的解也以概率密度函数的形式给出, 从而能更全面地刻画未知量.

统计学 (尤其是 Bayes 统计学) 的书籍亦不胜枚举, 它们与反问题关系密切却在反问题界鲜为人知. 反之, 诸如偏微分方程反问题和数学建模的内容在统计学界

目前尚未成为主流. 本书作为反问题和统计学之桥梁, 向双方介绍了其基本思想, 并结合两者发展出了一套 Bayes 统计反演的数学体系.

本书在理论上深入浅出, 在应用上由浅入深, 以大量的数值算例生动地佐证思想方法, 并阐释方法的实际应用价值. 本书前四章作为第一部分, 介绍了经典正则化、统计与非静态反演方法, 可作为从计算和统计角度研究反问题的入门材料; 后三章构成了本书的第二部分, 以案例分析的形式介绍了统计反演方法在理解经典方法和十几个现实问题中的具体应用.

本书全文翻译初稿由刘逸侃完成, 其后三位译者经过一年多的共同讨论和几度联合修改, 终于在 2017 年 8 月、2018 年 2 月访问东京大学期间定稿. 本书得到了国家基金项目 (批准号: 11331004、11471287、91534113、11871435) 和浙江省一流学科 (A 类) — 数学项目的支持, 在此深表感谢!

刘逸侃
东京大学大学院数理科学研究科
徐定华
浙江理工大学理学院 & 上海财经大学数学学院
程 晋
复旦大学数学科学学院 & 上海财经大学数学学院
2018 年 3 月

前　　言

本书面向应用数学专业的研究生, 同时也适合具有扎实数学功底的工科和物理专业学生阅读. 本书前四章可作为重点从计算和统计的方面关注反问题的入门教材. 另外, 第 3 章和第 4 章将讨论统计与非稳态反演方法, 适合已具备经典反演方法知识的学生阅读.

关于经典反问题方面的文献与教科书可谓浩如烟海. 从数值的观点来看, 在这类书籍关注的问题中, 测量误差被视为很小, 误差的性质已能确切获知. 然而在实际问题中, 误差通常不会很小, 且其性质在确定性的意义上也并非确切已知. 例如, 在经典文献中, 误差范数常被设为一已知实数, 而事实上误差范数是个随机变量, 其均值可能是已知的.

此外, 经典文献一般假设描述观测量的算子方程是精确已知的. 但同样, 通常在需要求基于实际测量的计算解时, 我们应考虑到, 数学模型本身也只是对实际现象的近似而已. 再者, 对问题的数值处理需将模型离散化, 这又会引入额外的误差. 于是, 测量值与观测模型的预测值之间的差异就不仅源于 “被加到测量值里的噪声”. 本书的主要课题之一就是对模拟所产生的误差进行统计分析.

在统计学 (尤其是 Bayes 统计学) 的文献中, 与反问题息息相关的亦不胜枚举. 这类文献在反问题领域鲜为人知, 因此本书的主要目的就在于向该领域引入统计学的概念. 对统计学家而言, 本书包含的新信息 (例如取样方法) 可能很少. 然而, 在偏微分方程以及相关的模拟误差分析等的基础上, 建立实际观测模型可能是有用的.

关于文献引用, 我们在第 1 — 6 章中主要引用了可作深入阅读的书籍, 但并不谈及这些课题的发展史. 第 7 章讨论我们先前的和一些最新的研究课题, 但同样不包含对其应用的评论. 这里我们主要引用原始文献以及包含改进与拓展的资料, 以此阐明统计方法的潜力.

第 5 — 7 章构成了本书的第二部分, 在其关注的问题中, 测量误差、无误差观测以及未知量都被视作模型, 它们本身可能包含不确定性. 举例来说, 一些观测模型基于偏微分方程和边值问题, 但部分边值数据可能本质上是未知的. 这样, 我们就尝试将这些边界数据模拟为随机变量, 既可将它们作为二级未知量来处理, 又可视作不确定性的进一步来源, 并计算其对观测模型与观测模型给出的预测之间差异的影响.

在举例时, 尤其是在第 7 章讨论非平凡问题时, 我们集中于我们早先开展的研究; 但我们也论述一些尚未发表, 或比原始文献更严密的课题.

我们避免在正文中引用文献, 以此来尽量增强本书的可读性. 每章 (除第 1 章) 都附有 "注释与评论" 一节, 其中给出了参考文献与可供深入阅读的材料, 以及对更高等课题的简短评论.

我们在此对我们的同事兼好友 Markku Lehtinen 表示感谢, 他数十年来始终倡导反问题的统计方法的研究, 并引导我们关注这一课题. 本书第 7 章的诸多成果是与我们现在和过去的研究生 (以及其他科学家) 共同完成的, 与他们共事可谓荣幸之至, 在此谨致谢忱. 这里我们仅提述那些通过改进其计算实现或其他方式对本书作出直接贡献的人: Ville Kolehmainen 博士 (7.2 节与 7.9 节), Arto Voutilainen 博士 (7.4 节), Aku Seppänen 先生 (7.5 节与 7.7 节), 以及 Jenni Heino 女士 (7.8 节). 我们也要衷心感谢 Daniela Calvetti 仔细阅读并点评了本书, 同时感谢上述各位阅读了本书的一些部分. 我们将对书中存在的疏漏承担全部责任.

本书的出版得到了芬兰学会与芬兰文理学会的经费支持, 在此表示诚挚谢意. 同时也要感谢 Kuopio 大学反问题课题组以及应用物理学系副主任 Ari Laaksonen 博士, 他在另一作者离开时接手并负责了后者的工作. 还要感谢新西兰 Auckland 大学的 Geoff Nicholls 博士与 Colin Fox 博士, 本书中许多新素材是作者在该校访问时构思的.

Jari Kaipio, Erkki Somersalo

2004 年 6 月于 Helsinki 和 Kuopio

目　录

第 1 章　反问题与对测量的诠释

顾名思义,“反问题”可以定义为直接问题或正问题的逆. 这种定义显然是空泛的,除非我们能定义“正问题”的概念. 通常,人们在对所关心的物理量进行间接观测时,就会遇到反问题. 让我们考虑一个例子:人们想知道温度. 温度本身是一个在统计物理学里定义的物理量,尽管它很常用且直观上易于理解,但它却不能被直接测得. 常见的温度计能够提供气温的信息,这依赖于以下事实:在一般条件下,水银等物质会随着温度的上升很有规律地膨胀. 这里的正向模型就是水银的体积关于温度的函数. 本例中的反问题是平凡的,以至于一般来说,根据测得的体积确定温度的问题并不被看成是一个独立的反问题. 而试图测量锅炉中的温度就是一个更具挑战性的反问题. 高温导致传统的温度计无能为力,这样我们就要采用更先进的手段,其中一种就是用超声波. 高温使得锅炉内的气体产生湍流,从而改变其声学性质,而这又反过来以回声反映出来. 现在的正向模型就包含了将湍流描述为温度和介质中声波传播的函数这一挑战性的问题,而通过测量声波来确定温度这一相应的反问题则更具挑战性.

自然规律通常可以表述为微分方程系统,这是 Newton, Leibniz 等留给后人的宝贵遗产. 在给定的点上,这些方程表现了函数及其导数在该位置对于物理条件的依赖性,从这个意义上说,它们是*局部的*. 这些规律的另一典型特征是*因果性*:后发情况取决于先发条件. 局部性与因果性是正向模型具有的典型特征,而恰恰相反,绝大多数反问题是*非局部且/或非因果的*. 在关于锅炉温度测量的例子中,外部测得的声波取决于整个内部的湍流,且由于信号速度是有限的,我们只可能重构测量之前一段时间的温度分布,即在计算上试图回溯以前的时间.

反问题的非局部性与非因果性很大程度上造成了其不稳定性. 为理解这一点,考虑物体内的热传导. 初始温度分布的微小变化被时间抹平,而最终时刻的温度分布几乎不变. 这样正问题就是稳定的,因为结果几乎不受初始数据变化的影响.

至于非因果性的方面,如果我们试图通过最终时刻的温度分布来估计初始温度分布,我们将会发现,至少在我们测量的精度范围之内,截然不同的初值条件可能产生同一个终值条件. 一方面,这一严重的问题要求我们对数据进行仔细分析;另一方面,我们必须纳入一切可能在测量之前获得的有关初始数据的信息. 作为本书的主题,*统计反演理论*通过在模型中恰当地纳入所有可以得到的信息来系统地求解反问题.

统计反演理论将反问题改述为通过 Bayes 统计学进行统计推断的问题, 所有的量在 Bayes 统计学中都被模拟为随机变量. 随机性反映了观测者对其数值的不确定性, 它蕴含在这些量的概率分布之中. 从统计反演理论的观点来看, 反问题的解是当所有可用信息都被纳入模型中时所关心的量的概率分布. 这一分布被称为*后验分布*, 它描述了这些量在进行测量之后的可信度.

不同于其他诸多的反问题教材, 本书并不关注分析性结论, 比如反问题解的唯一性或其先验稳定性问题. 但这并不意味着我们没有认识到这些结论的价值; 相反, 在分析除了实际测量外还需要哪些附加信息时, 我们认为唯一性与稳定性结论是十分有帮助的. 事实上, 在统计反演理论中, 设计涵盖所有先验信息的方法就是其中的一个巨大挑战.

反问题的教材中还有另一条路线, 它们强调*不适定问题*的数值解法, 并以正则化技巧为重点. 同样, 它们的观点亦与我们不同. 正则化技巧通常致力于在可用数据的基础上产生所关心的量的一个*合理估计*. 而在统计反演理论中, 反问题的解不是单个的估计, 而是一个能用来产生估计的概率分布. 但它给出的信息比单个估计更多: 它能产生截然不同的估计, 并评估它们的可靠性. 本书中第 2 章讨论了最常用的正则化方法, 这不仅是因为它们本身是很有用的工具, 更因为从 Bayes 统计学的观点诠释和分析这些方法能提供重要的信息. 我们相信, 这将有助于揭示这些方法是建立在何种隐含假设的基础之上的.

1.1 介绍性示例

本节我们通过典型的例子来说明前面讨论过的问题. 第一个例子是关于由反问题非因果性的本质引起的问题.

例 1.1 设有一根导热系数为 1 的单位长度细棒, 其两端温度固定为 0. 根据标准模型, 温度分布 $u(x,t)$ 应满足热方程

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t} = 0, \quad 0 < x < 1,\ t > 0,$$

以及边界条件

$$u(0,t) = u(1,t) = 0$$

和给定的初始条件

$$u(x,0) = u_0(x).$$

我们考虑如下的反问题: 给定时刻 $T > 0$ 的温度分布, 则初始温度分布如何?

我们首先将方程的解写成 Fourier 分量的形式:

$$u(x,t) = \sum_{n=1}^{\infty} c_n \, \mathrm{e}^{-(n\pi)^2 t} \sin n\pi x,$$

其中系数 c_n 为初始状态 u_0 的 Fourier 正弦系数, 即

$$u_0(x) = \sum_{n=1}^{\infty} c_n \sin n\pi x.$$

于是, 为确定 u_0, 我们只要利用终值数据解出系数 c_n 即可. 但假如我们有两个初始状态 $u_0^{(j)}$, $j=1,2$, 它们之间只相差一个高频成分, 即

$$u_0^{(1)}(x) - u_0^{(2)}(x) = c_N \sin N\pi x,$$

其中 N 很大. 这样, 相应的解在最终时刻就相差

$$u^{(1)}(x,T) - u^{(2)}(x,T) = c_N \,\mathrm{e}^{-(N\pi)^2 T} \sin N\pi x,$$

即由两个初始状态造成的终值数据的差别是指数级的小量; 因此, 任何关于高频成分的信息都将由于测量误差的存在而丧失. ◇

例 1.2 考虑由非均质性引起的时谐声波的散射. 在线性声学的框架内, 声压场 u 应满足波动方程

$$\Delta u + \frac{\omega^2}{c^2} u = 0 \quad 在\ \mathbb{R}^3\ 中, \tag{1.1}$$

其中 $\omega > 0$ 为**时谐依赖的**角频率, $c = c(x)$ 为传播速度. 假设在一个有界集 $D \subset \mathbb{R}^3$ 之外, $c = c_0 =$ 常数, 则改记

$$\frac{\omega^2}{c(x)^2} = k^2(1 + q(x)),$$

其中 $k = \omega/c_0$ 为波数, q 是一个定义为

$$q(x) = \frac{c_0^2}{c(x)^2} - 1$$

的具有紧支集的扰动. 假如我们沿着方向 $\omega \in S^2$ 发射一束平面波 u_0, 则整个场就可分解为

$$u(x) = u_0(x) + u_{\mathrm{sc}}(x) = \mathrm{e}^{\mathrm{i}\,k\omega\cdot x} + u_{\mathrm{sc}}(x),$$

其中散射场 u_{sc} 在无穷远处满足 Sommerfeld 辐射条件

$$\lim_{r\to\infty} r\left(\frac{\partial u_{\mathrm{sc}}}{\partial r} - \mathrm{i}\,k u_{\mathrm{sc}}\right) = 0, \quad r = |x|, \tag{1.2}$$

而场 u 满足 Lippmann-Schwinger 积分方程

$$u(x) = u_0(x) - \frac{k^2}{4\pi}\int_D \frac{\mathrm{e}^{\mathrm{i}\,k|x-y|}}{|x-y|} q(y)u(y)\,\mathrm{d}y. \tag{1.3}$$

将积分核关于 $1/r$ 作 Taylor 级数展开, 可知散射部分具有渐近形式

$$u_{\mathrm{sc}} = \frac{\mathrm{e}^{\mathrm{i}\,kr}}{4\pi r}\left(u_\infty(\widehat{x}) + \mathcal{O}\left(\frac{1}{r}\right)\right), \quad \widehat{x} = \frac{x}{r},$$

其中的函数 u_∞ 称为*远场模式*, 可由下式得到

$$u_\infty(\widehat{x}) = -k^2 \int_D \mathrm{e}^{-\mathrm{i}\,\widehat{x}\cdot y} q(y) u(y)\,\mathrm{d}y. \tag{1.4}$$

正散射问题就是在波速 c 已知的情况下确定声压场 u.

逆散射问题则希望通过不同入射平面波方向的远场模式的信息来确定未知的波速.

我们可以由此发现正问题与反问题的根本区别. 正问题需要求解满足辐射条件 (1.2) 的线性微分方程 (1.1), 或等价地求解 Lippmann-Schwinger 方程 (1.3), 它们都是线性问题. 而反过来说, 由于等式 (1.4) 中的 u 依赖于 q, 反问题却是强非线性的. 对这一问题可解性研究和数值求解需要相当高深的技巧. ◇

1.2 反演过失

本书我们将始终使用*反演过失*[①] (inverse crime) 这一术语. 反演过失意味着数值方法包含某些特性, 它们能有效地使反问题的不适定性降到实际问题本身具有的不适定性以下, 从而带来不切实际的乐观结果. 反演过失可简洁地概括为*将模型与现实视为同物*, 即研究者相信计算模型是精确的. 在实际中, 反演过失在这些情况下发生:

1. 数值上产生的模拟数据出自用来反演数据的同一个模型;
2. 用同一个离散格式进行数值模拟和反演.

整本书我们都会避免这些明显的反演过失. 此外我们还将说明, 统计反演理论能帮助我们分析模拟误差的影响. 我们将举例说明反演过失在模拟实例中会引起多大的差别, 以及更重要地, 正确模拟统计误差如何能有效地解决与离散化有关的问题.

---

① 据作者所知, 这一概念是由 Rainer Kress 在他的一个反问题的综述报告中引入的.

第 2 章　经典正则化方法

在本章中, 我们回顾在处理不适定反问题时最常使用的一些方法, 这些方法被称为正则化方法. 虽然本书并不侧重于正则化方法的技巧, 但理解这些方法的原理与工作机制仍是很重要的. 后面我们还将从统计学的角度分析这些方法, 这也是本书的主题之一.

2.1　绪论: Fredholm 方程

为了说明正则化的基本思想, 我们考虑一个简单的线性反问题. 依照惯例, 本章中的讨论均建立在 Hilbert 空间之上. 有关一些泛函分析结论的简要回顾可参见本书的附录 A.

设 H_1 与 H_2 为两个可分的有限维或无限维 Hilbert 空间, $A: H_1 \to H_2$ 为一个紧算子. 首先考虑如下问题: 给定 $y \in H_2$, 求满足方程

$$Ax = y \tag{2.1}$$

的解 $x \in H_1$. 我们称该方程为第一类 Fredholm 方程. 因为

1. 方程的解存在, 当且仅当 $y \in \mathrm{Ran}(A)$, 且
2. 方程的解唯一, 当且仅当 $\mathrm{Ker}(A) = \{0\}$,

所以以上两个条件必须同时满足, 才能保证问题具有唯一的解. 而从实际的角度来看, 求出一个有用的解还有第三个障碍. 向量 y 通常代表实测数据, 因此它会带有误差, 即我们得到的并非精确方程（2.1）, 而是近似方程

$$Ax \approx y.$$

众所周知, 即便 A 的逆存在, 它也可能是不连续的, 除非空间 $H_j (j = 1, 2)$ 都是有限维的. 于是, y 带有的微小误差可能会造成 x 的误差任意大.

例 2.1　反卷积问题是一个经典的不适定反问题. 设 $H_1 = H_2 = L^2(\mathbb{R})$, 并定义

$$A: L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad (Af)(x) = \phi * f(x) = \int_{-\infty}^{\infty} \phi(x-y) f(y)\,\mathrm{d}y,$$

其中 ϕ 为 Gauss 卷积核

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}.$$