

数值计算方法在高校数学统计问题中的应用探索

张迅¹ 蒙湘林²

玉林师范学院, 广西 玉林 537000

摘要: 在现代高校教育与科研活动中, 数学统计面临日益复杂且庞大的数据分析需求。面对多维度、多类型的数据, 传统的解析式或简单统计方法往往难以满足精度与效率并举的要求。数值计算方法凭借其灵活的算法设计与可控的误差分析, 成为高校数学统计问题的重要技术支撑。本文立足高校教学与科研实际环境, 系统探讨了数值计算方法与数学统计各方面的交互作用, 并对常见应用案例进行解析, 旨在为高校数学统计教育改革和科研效率提升提供可行的思路与实践路径。

关键词: 高校数学统计; 数值计算方法; 误差分析; 迭代算法; 教学改革

Exploration on the Application of Numerical Calculation Methods in Mathematical Statistics Problems in Colleges and Universities

Zhang,Xun¹ Meng,Xianglin²

Yulin Normal University, Yulin, Guangxi, 537000, China

Abstract: In modern college education and scientific research activities, mathematical statistics faces increasingly complex and huge data analysis requirements. Facing multi - dimensional and multi-type data, traditional analytical formulas or simple statistical methods often cannot meet the requirements of both accuracy and efficiency. Numerical calculation methods, with their flexible algorithm design and controllable error analysis, have become an important technical support for mathematical statistics problems in colleges and universities. Based on the actual environment of college teaching and scientific research, this paper systematically explores the interaction between numerical calculation methods and various aspects of mathematical statistics, and analyzes common application cases, aiming to provide feasible ideas and practical paths for the educational reform of mathematical statistics in colleges and universities and the improvement of scientific research efficiency.

Keywords: Mathematical statistics in colleges and universities; Numerical calculation methods; Error analysis; Iterative algorithms; Teaching reform

DOI:10.62639/sspfed20.20250103

引言

随着社会对数据分析与统计决策依赖程度的不断提升, 高校数学统计教育所面临的挑战与日俱增。不论在学术研究领域, 还是在各类具体应用场景, 海量且多元化的数据都需要快速且准确的处理, 这使得对数值计算方法的需求更为迫切。数值计算方法之所以能在高校数学统计中发挥愈发重要的作用, 根源在于它不仅能够在缺乏解析解或者无法直接进行精确推导时, 为问题提供近似可行的解, 而且可以通过算法的迭代过程, 对误差进行有效的控制与修正, 从而满足高精度要求。这一特点在现如今大数据、人工智能和云计算等科技蓬勃发展的背景下尤为显著。

与此同时, 高校数学教学本身也在进行转型, 从过去偏重纯理论知识与公式记忆的模式逐渐向“理论—方法—实践”三位一体的协同培养模式过渡。数值计算方法因其对编程实现、算法思路以及误差估计等环节有明确的要求, 成为培养学生综合能力的重要抓手^[1]。然而, 在具体教学和科研项目

中, 数值方法与传统统计方法的融合仍存在一定障碍, 很多学生因为在高等数学等课程中只关注公式的运用, 忽视了数值算法的推导和思想, 也对如何在软件工具(如 Matlab、R 语言或 Python)中实现算法缺乏直观的认识。基于此, 本文将结合具体的高校数学统计需求, 系统梳理数值计算方法的理论基础与实际应用, 以期对高校数学统计教学和科研工作具有一定的启示作用。

一、高校数学统计与数值计算方法的关系

(一) 理论与应用的需求差异

传统的高校数学统计课程更多关注统计原理、参数估计、假设检验等理论框架, 其优势在于能够系统地训练学生的抽象思维与推理能力。但是, 在面对真实世界中的复杂问题时, 往往会遇到高维度非线性模型、分布未知且数据噪声较大的情形, 很多理论上的经典公式或许难以直接适配。此时, 数值计算方法的优势凸显出来: 通过设计适当的算法, 不需要过多依赖封闭形式的解析解, 也能够有效控制

(稿件编号: FED-25-3-1019)

作者简介: 张迅 (2002-), 女, 汉族, 广西钦州市, 本科, 研究方向: 市场调研与消费者行为分析, 模式识别及算法。

蒙湘林 (2004-), 女, 汉族, 广西贵港, 本科, 研究方向: 应用统计。

基金项目: 2023 年大学生创新创业项目: “基于精准用户画像的国产手机智能推荐平台” (S202310606180X)。

误差的前提下给出满意的结果。

值得注意的是, 数值计算方法并非对任何问题都能“自动”给出答案。它需要基于对问题本质的深入理解来选择合适的算法思路。例如, 在处理某些回归分析或时间序列分析时, 如果数学模型的假设与实际情况偏差过大, 纯粹依赖数值算法往往会导致“垃圾进, 垃圾出”(GIGO)的后果^[2]。因而, 高校数学统计教学在引入数值计算方法时, 需注重对学生科研思维的培养, 既要让学生理解理论模型的适用条件, 也要掌握数值方法在实践中的具体实现与局限。两者相互支持、相辅相成, 才能提高对问题的整体把握与解决能力。

(二) 数值计算方法与现代科技工具

现代社会日新月异的科技发展, 为数值计算方法与统计分析的结合提供了强力支撑。早在二十世纪, 计算机的出现已经大幅度降低了数值方法在迭代运算上的成本, 让很多原本需要纸笔耗时演算的大规模计算变得可行。此后, 诸如 Matlab、Python、R 语言等一系列计算机工具的日益完善, 更是推动数值方法在高校科研与教学中快速普及。

然而, 随之而来的是同样需要被重视的问题: 过度依赖软件工具, 可能导致一些师生忽视了算法背后的数学原理。他们往往只会调用软件包中的现成函数, 却对如何选择、比较不同的数值方法、如何判断算法的收敛速度以及如何评估结果的可靠性等问题缺乏系统的认识。一旦遇到软件无法直接处理的定制化问题, 或需要深入优化算法时, 就会暴露理论基础薄弱的短板。因此, 高校有必要在数学统计课程中嵌入数值计算方法的原理和推导过程, 加强对学生批判性思维和算法设计能力的锻炼, 让其在实践操作中既能灵活使用软件工具, 也能通过理论思维发现和解决软件功能所不能直接覆盖的难点。

二、高校数学统计问题中的数值计算核心内容

(一) 插值与逼近方法

在统计学习和数据分析中, 插值与逼近是非常常见也十分重要的环节。高校学生常常需要在已知部分采样点或观测数据的前提下, 估计出未知区间内的函数值或趋势关系。最常见的莫过于多项式插值: 如果已经知道一些离散点, 就可以通过 Lagrange 插值、多项式拟合或样条插值等方法, 找到一个尽可能逼近真实数据分布的函数表达式。

然而, 多项式插值在维度较高或样本量较大的情况下, 计算量会急剧增长, 而且还可能出现振荡现象(如 Runge 现象)。此时, 学生若只停留在高校数学统计的“理论”层面, 往往缺乏对收敛特性的深入了解, 也不容易意识到数据分布对插值精度的影响。借助数值计算方法的分析思想, 可以引导学生理解为什么在一定区间内, 多项式插值可能变得不稳定; 为何分段插值或样条插值往往更有利于降低误差。对于逼近理论常用的“最佳平方逼近”或 Fourier 级数展开, 也能够让学生体会到“理论结果如何在实际数据下衔接与运用”的问题, 从而在具体问题求解时更具针对性。

(二) 误差分析与收敛性

当将数值计算方法运用于高校数学统计时, 误差分析与收敛性研究始终是绕不开的核心主题。统计推断中需要估计参数, 或要对模型进行假设检验, 都需要控制好结果的偏差和方差。如果采取数值方法, 则需保证算法在有限步内能以较快的速度收敛到理想解附近, 并保证最终结果的误差不超过预先设定的阈值。

例如, 在求解回归分析中某些非线性方程的最小二乘解时, Newton 迭代法、牛顿-科特斯法或共轭梯度法都可能被采用。这些方法各有优缺点: Newton 迭代法在初值选择合适时收敛速度快, 但若初值选择不当可能发散; 共轭梯度法适合大型稀疏系统且内存消耗相对低^[3]。对于高校教学而言, 不应只是向学生呈现这些方法的公式, 还应鼓励他们通过案例编程实践来验证不同迭代算法的速度与稳定性, 并通过数值试验理解何为收敛阶、何为截断误差、何为舍入误差及其成因。唯有对误差来源有了系统认识, 才能在面对真实科研与应用问题时, 有意识地采取措施减小或控制误差。

(三) 线性方程组与矩阵运算

在多元统计分析、主成分分析以及机器学习中的大量算法, 都需要高效求解线性方程组或对大型矩阵进行分解。此类问题虽然在高校数学课程中并不陌生, 但由于对规模较大的系统进行手动计算几乎不可能, 所以多数学生往往只在理论层面学过 Cramer 法则、线性变换与特征值分解等知识, 却缺乏对大型系统数值求解过程中的精度和效率的实感。

数值计算方法为此提供了丰富的手段, 包括 LU 分解、QR 分解、Cholesky 分解以及 Jacobi 迭代、Gauss-Seidel 迭代等方法。这些方法能够分别适用于不同结构和稀疏度的矩阵, 从而在教学与科研中发挥巨大作用。特别是在深度学习等前沿方向, 矩阵规模可达数百万维, 效率与精度的权衡极为关键。在教学中, 若能让学生编写或调用现有数值库的矩阵分解函数, 并结合实际数据来对结果进行检验与误差分析, 往往会比单纯讲授理论更具说服力和实用价值。

三、数值计算技术在高校数学统计教学中的应用

(一) 教学改革思路

在课程设计上, 可将数值计算方法贯穿于高校数学统计的各个知识模块。例如, 在学完线性统计模型后, 可以安排一个专题实验, 让学生基于真实或仿真的高维数据集, 尝试利用最小二乘法或岭回归等方法进行拟合, 并在过程中对不同迭代算法进行测试与对比。这样不仅能够使学生熟悉数值解的过程, 而且可以更加直观地理解统计模型的鲁棒性与稳定性。

另外, 结合现代科技手段是改革重点之一。随着多媒体教学和在线教育平台的普及, 授课教师可以使用交互式 Jupyter 笔记本或 Matlab Live 脚本等方式实时演示数值算法的推导过程及计算结果, 这大大缩短了学生对抽象公式和实际数据的认知鸿沟

[4]。通过可视化, 学生能够更快理解何为收敛曲线、何为迭代残差以及如何通过微调参数来改善算法性能。如此这般, 数值计算方法与统计分析知识能够更自然、更高效地融合。

(二) 结合科研项目的实践教学

科研与教学若能有效结合, 将极大提高学生的学习和主动探究能力。很多高校科研项目都涉及数学统计问题, 例如对实验室检测数据的统计建模、对教育测评数据的多维度分析、对校企合作项目中的生产工艺优化以及对自然科学、工程科学问题的大规模数据挖掘等, 都可以作为实训或课程设计的素材。

在这些项目中, 单纯的统计学理论往往不足以支撑数据处理的全部流程。若问题涉及非线性最优化、复杂动态模型或高频交易数据之类的应用, 就要借助数值计算方法来迭代求解。在此过程中, 学生需要学习如何为复杂模型选定合适的算法、如何评估算法收敛特性以及如何优化计算效率。经过多次实践后, 他们往往能意识到, “数值方法”绝不仅仅是一张公式推导的集合, 而是解决实际难题的“武器库”。学生也能更清晰地看见算法思路的演变过程, 培养耐心与细致严谨的科研态度。

四、高校数学统计中常见数值计算案例解析

(一) 回归分析与最小二乘拟合

在回归分析中, 对参数的估计通常可以转化为一类求最小化问题。例如, 线性回归的普通最小二乘法, 就是试图让残差平方和达到最小。假使问题规模不大且矩阵条件良好, 使用高职数学中学到的矩阵求逆或正交分解方法即可快速得到解析解; 但在高维或含有较多噪声、缺失值的情形下, 往往需要利用数值迭代算法来避免病态或奇异矩阵导致的不稳定性。

当学生面对高维数据, 且使用迭代算法(如梯度下降或随机梯度下降)进行回归分析时, 必须懂得在迭代时如何选取学习率(步长), 如何判断算法收敛, 以及如何利用正则化手段控制模型的过拟合。这样才能保证回归结果的数值稳定性与泛化能力。许多高校会以大作业或毕业设计的形式, 让学生对自己感兴趣的数据集进行回归建模, 以此检验其数值计算方法与统计理论的综合应用能力。

(二) 时间序列与差分方法

时间序列分析是高校数学统计课程的常见主题, 传统分析思路主要基于 ARMA、ARIMA 等模型, 关注平稳性、白噪声检验及自相关特性的检验。但随着实际应用中非平稳序列、非线性趋势以及外生冲击的关注度提高, 单纯依赖传统统计模型往往难以取得理想效果。此时, 可以将数值计算中的差分思想或迭代优化思想引入时间序列建模, 如在模型残差或状态方程的估计中, 引入数值求解策略, 实现对复杂序列的拟合与预测。

例如, 在需要同时考虑季节性和突发性事件的混合模型时, 解析方法很难得到闭合解, 但数值迭代方法可以逐步逼近最优参数集。特别是一些非线性预测场景下(如宏观经济指标预测、金融市场波动率建模), 常常利用数值方法对目标函数不断微

调, 更新模型结构与参数, 从而取得较高精度的预测结果[5]。通过这类案例, 可以帮助学生体会到数值计算和统计理论结合的威力: 前者解决求解困难的瓶颈, 后者提供模型设定与解释的框架。

(三) 多元统计与主成分分析

多元统计分析中, 主成分分析(PCA)作为一种降维与特征提取的方法, 被广泛应用于图像处理、基因数据分析、社会调查问卷维度精简等诸多领域。其核心在于对协方差矩阵进行特征值分解, 从而找出最大方差方向对应的特征向量。理论上, 特征值分解在小规模情况下可以直接基于高职数学中的线性代数知识来操作; 但对于大规模数据集, 这种方法往往需要求解规模巨大的线性方程组与高维矩阵分解, 计算量极其庞大。

此时, 就需要数值线性代数中的迭代方法, 如 Lanczos 方法或 Jacobi-Davidson 方法等, 以在较低的存储负担下逼近矩阵的主要特征值与特征向量。在高校教学实践中, 可以让学生对如图像分辨率降低、经济指标综合评价等真实案例进行 PCA, 并观察不同分解算法在运算时的稳定性与时间开销, 以及主成分的解释度, 从而掌握 PCA 在理论与数值实现之间的差异与联系。

五、结论与展望

数值计算方法与高校数学统计天然互补且协同发展。从理论上, 它为无法解析求解或过于复杂的问题提供了高效的近似解; 从实践上, 它不仅提升了统计分析的精度, 还丰富了高校教学内容, 为学生带来多元化的学习体验。在信息化与智能化背景下, 将数值计算方法深度融入高校数学统计已成为培养学生综合分析、创新研究及跨学科思维能力的重要途径。为实现更广泛且有效的应用, 需在课程设计、师资培训及实践项目等方面持续改进, 理论与实践结合尤为关键: 课程中需融入数值算法与误差分析, 实验中则需设计真实场景以提升学生动手能力。同时, 云计算、人工智能及大数据技术的成熟, 使得许多复杂问题逐步可解。未来, 高校数学统计需紧密结合学科前沿与社会需求, 将数值仿真与算法实践融入教学, 打破传统教学局限。数值计算不仅是辅助工具, 更是培养学生解决现实问题与开展创新研究的重要抓手, 助力其在复杂多变的时代中具备扎实理论与实践能力, 为社会发展提供更强有力的人才支持。

参考文献:

- [1] 陈倩云, 杨林锋. 大数据背景下高校财务统计分析工作存在的问题及其对策 [J]. 中国农业会计, 2024, 34 (10): 69-71.
- [2] 管策, 杨露豪, 郝玺龙. 基于 incoPat 的高校专利信息统计分析——以太原理工大学为例 [J]. 内蒙古科技与经济, 2024, (03): 47-51.
- [3] 胡贝贝, 崔振澜. 地方高校高等数学课程学习成绩影响因素的统计分析——以滁州学院为例 [J]. 西昌学院学报(自然科学版), 2023, 37 (02): 116-122.
- [4] 王联芳. 基于 Wigley 数学船型的兴波波形与尾迹数值计算 [J]. 舰船科学技术, 2021, 43 (08): 10-12.
- [5] 吴松林, 陈体英. 基于双星的干扰源定位数学模型及数值计算 [J]. 数学的实践与认识, 2020, 50 (16): 124-131.