

物理学中的 数据处理和误差分析

WULI KEXUE ZHONG DE (第三版)

SHUJU CHULI HE

WUCHA FENXI

[美] 菲利普·R. 贝文顿 D. 基思·鲁滨逊 著

夏元复 何云 译

Philip R. Bevington

D. Keith Robinson



GUANGXI NORMAL UNIVERSITY PRESS

广西师范大学出版社

物理学中的 数据处理和误差分析

(第三版)

[美] 菲利普·R. 贝文顿 D. 基思·鲁滨逊 著

○ 夏元复 何 云 译

Data Reduction and

Error Analysis for the Physical Sciences (THIRD EDITION)



GUANGXI NORMAL UNIVERSITY PRESS

广西师范大学出版社

· 桂林 ·

Philip R. Bevington, D. Keith Robinson
DATA REDUCTION AND ERROR ANALYSIS FOR THE PHYSICAL SCIENCES
THIRD EDITION
ISBN 0-07-247227-8

Published by McGraw-Hill, a business unit of The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY 10020. Copyright © 2003, 1992, 1969 by The McGraw-Hill Companies, Inc. All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of The McGraw-Hill Companies, Inc., including, but not limited to, in any network or other electronic storage or transmission, or broadcast for distance learning.

Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

本书中文简体字翻译版由广西师范大学出版社和美国麦格劳-希尔教育(亚洲)出版公司合作出版。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有 McGraw-Hill 公司防伪标签,无标签者不得销售。

著作权合同登记号桂图登字:20-2005-086 号

图书在版编目(CIP)数据

物理科学中的数据处理和误差分析:第3版/(美)
菲利普等著;夏元复,何云译. —桂林:广西师范大学出版社,2006.12

ISBN 7-5633-6345-9

I. 物… II. ①菲… ②夏… ③何… III. ①物理学—数据处理②物理学—误差分析 IV. O4

中国版本图书馆 CIP 数据核字(2006)第 151560 号

广西师范大学出版社出版发行

(广西桂林市中华路 22 号 邮政编码: 541001)
网址: <http://www.bbtpress.com>

出版人: 肖启明

全国新华书店经销

广西师范大学印刷厂印刷

(广西桂林市临桂县金山路 168 号 邮政编码: 541100)

开本: 787 mm × 1 092 mm 1/16

印张: 17.5 字数: 448 千字

2006 年 12 月第 1 版 2006 年 12 月第 1 次印刷

印数: 0 001 ~ 5 000 册 定价: 35.00 元

如发现印装质量问题,影响阅读,请与印刷厂联系调换。

内容提要

现代的物理学工作者必须掌握数据处理和误差分析的方法,它是从事物理学工作者素质的重要标志之一。本书自 McGraw-Hill 高等教育出版社出版以来,就一直是西方发达国家广泛使用的著名教材。本书由 2003 年出版的该书第三版译出。前四章引入下列基本概念:测量的不确定性、误差分析和概率分布。第五章为实验数据的模拟分析,引入蒙特卡罗方法,用前几章例子给出的数据研究和评估实验结果的统计意义。第六章至第九章讨论最小二乘法,第十章提供了最大似然法的直接应用。第十一章讨论 χ^2 概率和置信区间。

本书的另一个特点是附有简单直观的电脑程序。这些程序可以从出版公司网站上直接下载。论述有足够的广度和深度,以及强调实用的特点,对大学本科生和研究生都适用,也能使物理学工作者为此受益。

作者简介

Philip R. Bevington (菲利普·R. 贝文顿) 是 Case Western Reserve 大学物理学教授, 已故。他 1954 年毕业于哈佛大学, 1960 年在 Duke 大学得到 Ph. D. 学位。在 Duke 大学任教五年后, 1963 年至 1968 年为斯坦福大学助理教授, 此后到 Case Western Reserve 大学。他曾用静电加速器作核结构物理学研究。在斯坦福大学工作期间, 他将计算机应用于核物理的领域, 并负责网络系统自动程序控制 (SCANS) 的开发。

D. Keith Robinson (D. 基思·鲁滨逊) 是俄亥俄州克利夫兰市 Case Western Reserve 大学物理学教授, 1954 年在加拿大 Dalhousie 大学获得物理学学士学位, 1960 年在牛津大学获得 Dr. Phil. 学位。1960 年至 1966 年为布洛克海文国家实验室成员, 此后进入 Case Western Reserve 大学。在实验粒子物理领域的研究有: 玻色子共振, 反质子-质子相互作用, 超子的辐射衰变。他是将电脑引入 Case Western Reserve 大学物理课程的主要贡献者之一。

第三版前言

在本书 1969 年第一版的前言中,已故作者 Philip R. Bevington恰当地阐明了他编写此书的目的:“在于引入物理科学研究中普遍使用的数据处理和误差分析方法,并以足够详细和足够的广度介绍它,以使该书对本科生和研究生都有用。论述从实用出发,包括用充分的推导来证明结果,并强调方法多于理论。”第三版沿用了他的原有意图,更新了一些能使用电脑的部分内容。

第一章至第四章介绍了测量的不确定性、误差分析和概率分布等概念,增加了关于统计性实验中概率的新内容。第五章为实验数据的模拟分析,介绍了蒙特卡罗方法,用前几章的例子所给出的数据,研究和评估实验结果的统计意义。在第六章至第九章中,最小二乘法被用于更复杂的问题中,由解析的直线拟合到要求迭代求解的非线性拟合。第十章介绍了最大似然法的直接应用,第十一章讨论 χ^2 概率、置信区间和相关系数。在各章末都附有丰富的练习,从简单的统计计算到诸如用最小二乘拟合以及蒙特卡罗计算完成的小题目。部分练习有简要参考答案。

本次版本保留了以前各版本的附录。附录 A 包含了各种不同的基本算法,但主要是用于在电脑上分析数据的数值方法。附录 B 中讨论矩阵和行列式。附录 C 提供统计函数的图表,增加了为计算概率可以借助的网址上的电脑程序。附录 D 列出了为准备合格图示的若干指南。附录 E 给出了正文中的电脑程序。

电脑程序

简单、直观的电脑程序是本书第一版的特点,本版保留了它,并用 Fortran 77 语言在附录 E 中给出。选用 Fortran 语言是因为它在数十种语言中是用得最持久的。在第二版中所用的 Pascal 语言已经不用了,而为 C++ 语言所替代。借助附录 E 开头部分的说明,学生们应当能阅读 Fortran 程序,遵循其逻辑而不需要专门的语言技巧。为了简化所列程序并明确它们的主要目标,删去了图解程序中的多数调用程序。

Fortran 语言或 C++ 语言程序以及简化构建蒙特卡罗生成、最小二乘拟合以及概率计算的完整程序可由 www.mhhe.com/bevington 网址(website)下载,该网址上的“Read Me”文件阐明程序的组织,并提供使用指南。

感谢

我首先感谢已故的菲利普·R. 贝文顿先生(1933—1980),他的原书是这些修正版的基础。我感谢 Case Western Reserve 大学物理系对本书的支持,感谢我的大学本科实验室学生们提供了一些例子和很多启发。我也感谢 Spencer Cotkin 和他在 McGraw-Hill 的同事们的鼓励。

我还感谢读者们,特别是以下审阅人的很多百益评述和帮助:Riverside 加州大学的 Jingsong Zhang;亚利桑那大学的 Gary Schmidt;Berkely 加州大学的 Herbert Strauss;Kinsville 德克萨斯 A & M 的 Daniel Suson.

最后,我想感谢我的夫人 Margi,感谢她对我的支持和耐心。

D. 基思·鲁滨逊

出版说明

众所周知,数据处理和误差分析在物理学研究中占有十分重要的地位,也是物理学工作者素质的重要表现。本书在欧美发达国家近年本科教学中起着重要的作用,是很多著名大学的教材或者教学参考书。第一版由著名的 McGraw-Hill 公司在 1969 年出版发行,该公司 1992 年又发行了第二版。

本书是该公司 2003 年出版发行的第三版的中译版。其中第一章至七章由夏元复翻译,第八章至十一章,以及附录由何云翻译。利用电脑进行数据处理和误差分析是本书的特色。随书附上用 Fortran 语言和 C++ 语言编写的电脑程序(自网址 www.mhhe.com/bevington 下载,也可由 www.bbtpress.gxnu.edu.cn/bevington.asp 下载),阐明了程序的组织和提供了使用指南。采用该书作教材的教师可向 McGraw-Hill 公司北京代表处联系索取教学课件资料,传真:(010)62638354,电子邮件:webmaster@mcgraw-hill.com.cn。中译版由广西师范大学出版社出版,何云参与了全书的校阅。

译文如有不完善之处,恳请广大专家读者指正。

夏元复 何云
2006 年 8 月

序

物理学,甚至任何一门自然科学,都是建立在实验或观测的基础上。实验和观测的直接结果都是数据。数据,无论直接用肉眼观测,还是用仪器测量,总是有误差的。要从数据得到有意义的物理结果,必须经过数据处理和误差分析。所以,任何一个实验物理工作者,都必须掌握数据处理和误差分析,才能准确地给出测量所得的物理结果。其实,一个理论物理工作者也应懂得数据处理和误差分析,否则不能正确理解实验给出的物理结果,也不能正确地将理论计算与实验结果进行恰当的比较。

特别是在大科学项目不断涌现的今天,比如大型加速器、大型望远镜、空间探测器(大型卫星项目)等物理实验或天文观测项目,实验或观测给出的都是大量数据构成的数据库,任何新的发现或者新的规律都是从处理和分析这种数据库而得到的。1989年发射升空的 COBE 卫星,观测宇宙微波背景辐射,收集了大量数据,经几年的分析处理,发现了微小的、只有 10 万分之一量级的各向异性。这个微小的各向异性却是人们所能直接观测的最远、也是最早的宇宙幼年时的照片。后来,2001 年发射升空的 WMAP 卫星,进一步提高了观测精度,得到了大量精确数据。第一年的观测数据经处理后于 2003 年发表。2006 年又发表了三年观测数据的处理结果。这里,数据处理和误差分析的工作量非常大,由此得到了许多重要结果:测定了宇宙的年龄、定出了宇宙物质的组成,确定了宇宙的空间几何性质。此项成果导致了 John Mather 和 G. Smoot 两人获得了 2006

年度的诺贝尔物理学奖。事实上,几乎所有的实验研究均要处理大量的数据,而且不少从事理论研究的人,也经常要充分利用观测数据来寻找新的规律。1985年,M. van der Klis 在处理低质量 X 射线双星的观测数据时,发现了准周期振荡(QPO)新现象,就是一例。

数据处理和误差分析既然如此重要,在培养物理学科的人才时,选择一本合适的教材是十分重要的。所幸由 Philip R. Bevington(菲利普·R. 贝文顿)和 D. Keith Robinson(D. 基思·鲁滨逊)撰写的《物理科学中的数据处理和误差分析》第一版在1969年由 McGraw-Hill 公司出版以来,一直被许多国家广泛采用,非常成功。1992年该公司又出版了它的第二版。在2003年出版的第三版中,D. Keith Robinson 又作了较大的修改。这本书讲述的概念比较清楚,实用性也很强,特别是教材中引入了不少的电脑程序,适合于现代实验工作的需求。它不仅是一本好教材,也可以作为实际工作的一本好手册。对于培养物理学科的本科生和研究生非常合适,对于直接从事实验研究的物理工作者也十分有用。

这本中译本是由南京大学夏元复教授(俄罗斯科学院外籍院士)和广西师范大学何云教授根据该书最新版(第三版)翻译的。夏元复长期从事 Mössbauer 谱学的实验和理论研究,并在培养研究生的过程中已经多年实际采用了这本书,何云也在他培养研究生的过程中使用这本书,效果均非常好。两位译者本人又都是这方面的专家,翻译仔细,语言流畅,是一本很好的译本。

相信这本书对于学习物理学的本科生和研究生,以及从事物理工作的教师和研究人
员,均将是一本有益的好书。

紫金山天文台研究员

南京大学教授

陆 垓

中国科学院院士

2006年11月28日

目 录

第三版前言

出版说明

第一章 测量中的不确定性

- 1.1 测量误差 /001
- 1.2 不确定性 /005
- 1.3 原始分布和取样分布 /006
- 1.4 分布的平均值和标准偏差 /010

第二章 概率分布

- 2.1 二项式分布 /015
- 2.2 泊松分布 /021
- 2.3 高斯分布和正态误差分布 /025
- 2.4 洛伦兹分布 /028

第三章 误差分析

- 3.1 仪器误差和统计误差 /034
- 3.2 误差传播 /037
- 3.3 具体的误差公式 /039
- 3.4 误差方程的应用 /044

第四章 平均值估计和误差计算

- 4.1 最小二乘法 /049
- 4.2 统计涨落 /057
- 4.3 概率检验 /060

4.4 分布的 χ^2 检验 /061**第五章 蒙特卡罗法**

5.1 引言 /071

5.2 随机数 /074

5.3 来自概率分布的随机数 /076

5.4 特定分布 /079

5.5 有效的蒙特卡罗生成 /089

第六章 直线的最小二乘拟合

6.1 因变量和自变量 /093

6.2 最小二乘法 /097

6.3 使 χ^2 最小 /099

6.4 误差估计 /101

6.5 最小二乘法的局限性 /104

6.6 其他拟合法 /105

第七章 多项式的最小二乘拟合

7.1 行列式解 /110

7.2 矩阵解 /115

7.3 独立参量 /120

7.4 非线性函数 /127

第八章 任意函数的最小二乘拟合

8.1 非线性拟合 /133

8.2 搜索参数空间 /138

8.3 格点搜索法 /141

8.4 梯度搜索法 /143

8.5 展开法 /146

8.6 麦夸特法 /151

8.7 评述 /153

第九章 拟合复合曲线

9.1 二次型本底上的洛伦兹峰 /158

9.2 面积的确定 /160

9.3 复合曲线 /163

第十章 最大似然法的直接应用

10.1 最大似然法简介 /169

10.2 计算机举例 /176

第十一章 拟合检验

11.1 拟合优度的 χ^2 检验 /182

11.2 线性相关系数 /185

11.3 多变量相关性 /188

11.4 F 检验 /191

11.5 置信区间 /195

11.6 蒙特卡罗检验 /199

附录 A 数值算法

A.1 多项式内插 /204

A.2 微积分基础:微分和积分 /208

A.3 数值微分和积分 /212

A.4 三次仿样函数 /214

A.5 非线性方程的根 /217

A.6 数据的平滑化 /220

附录 B 矩阵

B.1 行列式 /223

B.2 用行列式求解联立方程 /228

B.3 矩阵求逆 /230

附录 C 图和表格

C.1 高斯概率分布 /233

C.2 高斯分布的积分 /234

C.3 线性相关系数 /237

C.4 χ^2 分布 /238C.5 F 分布 /244C.6 学生的 t 分布 /245

附录 D 矩形图和曲线图

D.1 作曲线图 /251

D.2 参数的图解计算 /252

D.3 矩阵图和频率曲线 /254

D.4 绘图子程序 /256

附录 E 用 Fortran 语言的电脑子程序

程序列表(略)

参考书目 /260**部分练习题答案 /263****后记 /266**

第一章 测量中的不确定性

1.1 测量误差

科学研究已经确认了以下事实:最初实验所得的结果与所寻求的“真实”只是稍有相似。当重复实验,并逐渐改进测量技术和方法,测量结果才逐渐逼近答案,这个答案我们可以以一定的置信度接受,认为是对事件的可信赖的描述。我们有时感觉到自然界并不愿意轻易放弃它的秘密,如果我们不经过巨大的努力,实验一开始就注定会失败。不管什么理由,对所有物理实验都必须确信,误差和不确定性是存在的,并必定会通过改进实验技术和重复测量使之减小,而所得误差常常能用来估计我们的结果的真实性。

Webster 是这样定义误差的:“观测值(或计算值)与真实值之间的差别。”当然,通常我们不知道“真实”值,否则就不需要做实验了。然而,我们也许从以前的实验或者从理论预测大致知道它应当是怎样的。这样的近似可以作为导向,但我们常常必须从数据和实验条件本身,系统地确定我们的实验结果能有多大的置信度。

有一类误差我们是立即处理的:误差来自实验中或计算中的错误或疏忽。所幸的是,这些误差通常是明显的,或者是显然不正确的数据点,或者是不合理地接近估计值。它们被归入“不合理误差”,通常可以通过仔细地重复操作得到校正。更要考虑的是在测

量中随机涨落引起的不确定值以及系统误差,它们以大致明确的途径限制着结果的精密度和准确度。通常,将不确定值称为误差,而将估计它们的过程称为误差分析。

准确度和精密度的关系

区分**准确度**和**精密度(精度)**这两个术语是很重要的。实验的**准确度**量度实验结果接近真实值的程度。实验的**精密度**量度被确定的结果是否精细,但不管它是否与真实值相符。此外,精密度也是对实验结果可重复性的量度。准确度和精密度之间的区别如图 1.1 所示。其中各图中的直线表示因变量 y 和自变量 x 之间的预期关系。在这两个图中,数据点的分散反映了测量中的不确定性,即数据点上的误差线。由短的误差线说明测得的图 1.1(a)的数据精确度高,并与预期的 y 对 x 的变化符合得好,但显然不准确,为偏高直线。另外,由长的误差线说明图 1.1(b)中的数据点不精确,但它散布于预期分布的周围。

显然,对任何实验我们必须同时考虑准确度和精密度。如果我们已经知道结果是高度不准确的,得到的结果即使具有高精密度也是白费时间和精力。相反,如果精密度低,这个结果就不能被认为是十分准确的。一般说来,当标出实验结果的不确定性或误差时,指的是被确定结果的**精密度**。**绝对精度**指结果中不确定性大小与结果值用相同单位表示。**相对精度**指不确定性以结果数值的百分数表示。

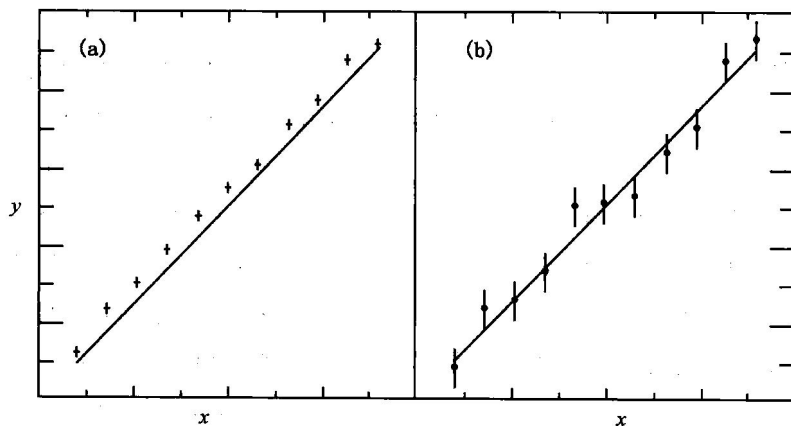


图 1.1 图示精密度和准确度间的差异。(a)精密但不准确的数据。

(b)准确但不精密的数据。真实值用直线代表。

系统误差

如同我们已经定义的,一般说来,实验的**准确度**取决于我们能在多大程度上控制或补

偿系统误差。系统误差引起测量的结果与“真实”值之间存在可重复的偏差。这类误差不易被发觉,而且不易用统计分析来研究。它们会来自设备不合格的标定,或者来自一些观测者的读数偏差。必须通过分析实验条件和实验技术来估计。设计实验的主要内容应当是了解和减少系统误差的根源。

例 1.1 某学生用钢制米尺测量桌面,他测量所得桌子长度的平均值为 (1.982 ± 0.001) m。米尺是在 25°C 时校正的,膨胀系数是 $0.005^{\circ}\text{C}^{-1}$ 。测量是在 20°C 室温时进行的,真实值应当整体变小。为了校正该效应,他将结果乘以 $1 + 0.000\ 5 \times (20 - 25) = 0.997\ 5$,于是重新确定长度为 1.977 m。

当该学生重复实验时,他发现第二个系统误差,即他读米尺的技术有误,以致读数总有偏差。由实验他确定读数缩短了 2 mm。从而经校正后的结果为 1.979 m。

在本例中,得出的第一个结果有很高的精度,精度约 $1/2\ 000$ 。该结果的校正正是通过最佳估算来补偿第一个结果的已知偏差来源实现的。这些校正完全不改进精密度,事实上还使其变差,因为校正本身只是试图算出恰当的校正值。在引述其最后结果前,该学生必须重新核实其误差分析,并计入由这些校正引入的所有其他误差。

随机误差

一个实验的精度取决于能在多大程度上克服随机误差。观测值的上下波动造成逐次实验所得结果有差异,因此要求重复实验以得到精确的结果。一个给定的准确度意味着相应的精度,并因此也在一定程度上与随机误差有关。

减小随机误差的本质是改进实验方法和优化实验技术,以及简单地重复实验。如果随机误差来自测量仪器的不确定度,可以通过用更可靠和更精密的测量仪器来减小误差。如果随机误差来自测量次数的有限造成的统计波动,可以通过更多的测量次数来减小误差。这些改进都具有实际限度。在例 1.1 对桌子长度的测量中,该学生可以试用放大镜阅读标度以改进其测量精度,或者可以通过重复测量若干次来减小测量中的统计波动。无论哪一种情形都无助于将随机误差减小到远小于系统误差,系统误差诸如因米尺的校准或者误读标度引起。在设计实验过程中必须着重考虑限制系统误差。