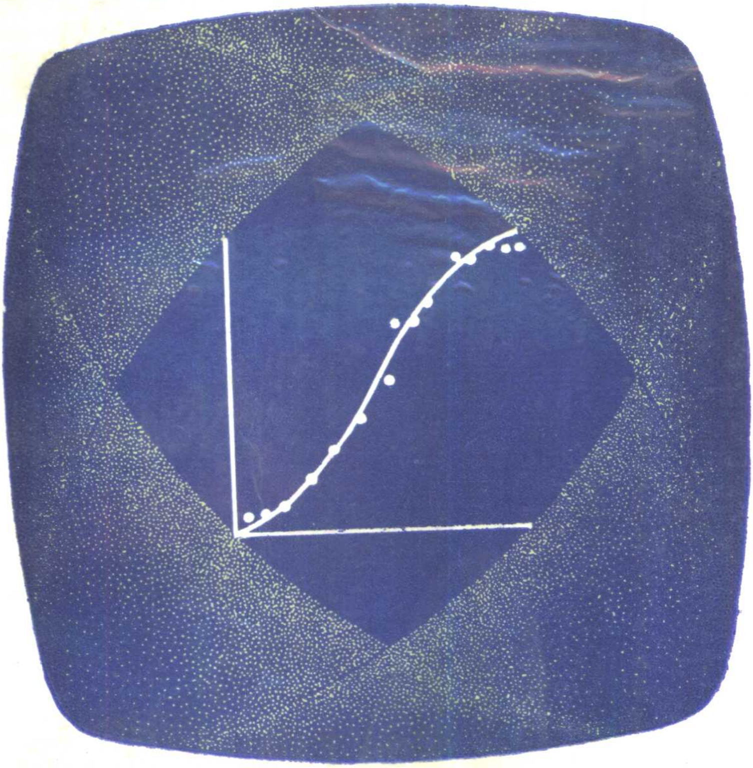
D·A·Ratkowsky 著



洪再吉　韦博成　吴诚鸥等　译

高金衡　校

非线性回归模型

统一的实用方法

南京大学出版社

非线性回归模型

统一的实用方法

D. A. Ratkowsky　著

洪再吉　韦博成　吴诚鸥　等译

高金衡校

南京大学出版社

1986·南京

内 容 提 要

本书主要叙述非线性回归的常用数学模型和方法，讨论有关非线性回归的估计及其性质等问题，书中除介绍必要的基本理论外，还载有大量的结合实际问题提出的模型和许多应用实例，书末附有 FORTRAN计算机程序，便于效用。本书可作为非线性回归分析的入门书，对广大实际工作者，特别是在农业、生物、化工、电子、气象以及经济等部门中的数理统计工作者，有很大的应用和参考价值。

本书对象为需要应用数理统计的科技人员、教师、研究生以及高年级大学生等。

David A. Ratkowsky
NONLINEAR REGRESSION MODELING
A Unified Practical Approach
Marcel Dekker, 1983

非 线 性 回 归 模 型

统一的实用方法

D. A. Ratkowsky 著

洪再吉 韦博成 吴诚鸥 等译

高金衡 校

南京大学出版社出版

（南京大学校内）

江苏省新华书店发行 国营阜宁印刷厂印刷

开本 787×1092 1/32 印张 10.25 字数 230千

1986年9月第1版 1986年9月第1次印刷

印数 1—3,000

统一书号：13336·022 定价：2.50元

翻译说明

《非线性回归模型》这本书是由澳大利亚联邦科学和工业组织(CSIRO)数学统计部David A. Ratkowsky撰写的，由美国Marcel Dekker公司作为数理统计丛书中的第48本于1983年出版的。该书主要叙述非线性回归的常用数学模型和方法，讨论有关非线性回归的估计及其性质等问题；书中除介绍必要的基本理论外，还载有大量的结合实际问题提出的模型和许多应用实例；书末附有FORTRAN计算机程序，便于效用。我们相信，本书的翻译出版对我国非线性回归的应用和发展是有益的。由于D. M. Bates和D. G. Watts的论文“非线性性相对曲率的度量”是本书引用的最主要论文，所以我们也将它的主要部分译出，附在书末，以方便读者参考。

在南京工学院应用概率统计教研室高金衡教授的主持下，南京工学院、南京大学、南京气象学院和工程兵工程学院等院校的一些有关教师联合组织了小型数理统计讨论班，本书就是在讨论班基础上组织翻译出版的。

本书的第1，3两章由陈华钧译出，第2章由韦博成译出，第4章和书末的练习解答以及所附的“非线性性相对曲率的度量”由洪再吉译出，第5，7两章由鲁国斌译出，第6章由吴诚鸥译出，第8章由赵星译出，第9章由孔凡辉译出，书末附录由范博舫和史建清译出。另外，洪再吉负责全

书翻译的统一工作。全书译稿由高金衡教授校对。

武汉大学张尧庭教授对本书的翻译给予热情的鼓励，并在百忙中为本书翻译写了前言。南京工学院陶永德教授对本书的翻译十分关心。南京大学出版社为本书翻译的出版给予热情的支持。在此一并向他们表示衷心的感谢。

由于译者水平所限，误译和不妥之处，敬请读者批评指正。

译　者

1985年11月

前　言

在众多的统计方法中，线性回归的方法是应用范围最广，收效甚多的一种。但是，搞过实际工作的同志都体会到，有些问题实质上是非线性的，因为没有处理非线性问题的方法，只能用近似的线性方法，往往效果是不理想的.在理论上，非线性问题比线性问题难处理得多，即使是估计参数，要得一个简单的表达式已不容易，往往只能讨论大样本的情形。Ratkowsky的这本书是在这一方面着重讨论实用的处理方法的第一本专门著作，能及早翻译出版这样一本书，不仅对实际工作者有很好的参考价值，还将会引起理论工作者的兴趣，推动这一专题的理论研究。本书的作者在书的序言中已经指出，对于非线性回归“拟合”的好坏如何分析，这本书是不讨论的。这对于从数据出发如何建立模型来说，无疑是一个重要的问题。希望这本书的出版能推动这一问题的研究。

用微分几何的方法来处理统计中的问题，是七十年代末期引起国际上统计界关注的一个方向，经过许多人的努力，得到了不少重要的理论成果，但是在实用的方法上，没有重大的突破。本书在这一方面是有特点的，叙述了非线性程度的一个度量，这一度量从微分几何看来是很自然的，从这个意义上说，微分几何方法已经给统计带来了新的方法.当然，这个方法的效果如何，还需从理论和实践这两方面做很多工

作后才能判断，但从本书的例和讨论来看，是很前有途的。

我自己对非线性回归很不熟悉，也想通过这本书来学习一些有用的知识，衷心希望这本书的翻译、出版能把非线性回归方向推前一步，几年之后，能形成一个高潮。

张尧庭

1986年元月

序　言

在线性回归模型中出现的参数都是线性的，而在非线性回归模型中至少出现一个参数是非线性的。本书致力于研究非线性回归模型估计的性态，这一论题的研究在很大程度上仍然处于初始阶段。尽管非线性回归模型有大量的人使用，寻求非线性回归模型的参数的最小二乘估计量的方法也有大量的文献，但仅在近二十年才对这些估计量的性质作一些认真的研究。即使如此，在这二十年期间中有关这一主题论文的数量仍然非常的少。在估计的性质方面，非线性回归模型与线性回归模型的重大差别在于：在给定独立同分布正态随机项的通常假定后，线性模型产生无偏的、正态分布的、极小方差估计，而非线性回归模型一般地只当样本大小很大时才是那样。本书的目的是探讨样本大小与在农业研究、生物、工程以及其它应用学科的科学家的实践中实际得到的样本大小相近的问题中估计量的性质。

研究模型者给一组数据拟合一个非线性回归模型时可能有以下三个目的之一。他们的兴趣可能在于：（1）仅仅为了表示的目的，对给定的数据求得一个“好的拟合”，（2）对回归变量X的给定值，预测响应变量Y的值，以及（3）基于参数估计的解释作出推断。本书只涉及第二和第三个目的，第一个目的不在讨论范围之内。在非线性回归模型中，Y的预测值将是有偏差的，偏差的程度依赖于通常称为模型数据集组合的“固有非线性”（intrinsic nonlinearity）的量，参数估计的解释也包含在另一个通常称为“参数效应非线性”(parameter-effects nonlinearity）的量中．即使固有非线性是低到可接受的程度，但是高的参数效应非线性可能使得利用标准算法例如高斯-牛顿（Gauss-Newton）法来得到收敛于参数的最小二乘估计（least-squares estimates）这件事变得困难了．贯串全书的论点是：假定固有非线性是可接受的低，则希望寻找具有低参数效应非线性的模型，而这样的模型常常可以用适当的重新参数化来得到．一个适当的重新参数化的模型的性能就象一个线性模型一样，这时高斯-牛顿法将很快收敛于最小二乘估计，这样就使得为得到参数估计而发展起来的各种复杂方法都不需要了．

假定本书的读者已具有统计方法和计算的基础以及有例如Draper和Smith（1981）与Gunst和Mason（1980）的标准教科书中所介绍的那种水平的线性回归分析的理论和实践的知识．那些接触过应用统计或生物统计的各实际学科的科学家们也应该能够使用本书中提出的方法*．

D．A．Ratkowsky

*译者注：以下感谢的话，从略。

目　　录

3 产量-密度模型

4 S形生长模型

5 渐近回归模型

1

回归模型介绍

1.1 线性回归模型

术语“线性回归模型”往往有两种不同的意义。首先它指的是用于两变量之间的直线关系。其次它指的是一个模型，在模型中要估计的参数是线性地出现的。本书始终用后一用法。下面是一些线性回归模型的例子：

$$Y=\alpha+\beta X+\varepsilon \tag{1.1}$$

$$Y=\alpha+\beta X+\gamma X^2+\varepsilon \tag{1.2}$$

$$Y=\alpha+\beta_1X_1+\beta_2X_2+\varepsilon \tag{1.3}$$

这里希腊字母α，β，β_1，β_2和γ表示未知参数，对于给定的模型数据集组合，参数被认为是常量。X，X_1和X_2是变量（常称为回归量，预报因子或独立变量），它们可以表示实

验装置，预定条件，或者是不可控观察值但假定测量没有误差。一般地说它们不是随机变量。响应变量Y也称为因变量，与回归线给出的期望值（平均值）的偏离为ε，它是一个不可观察的随机误差项，它的值是未知的但假定其均值为零。

模型（1.1）可以看作为“直线回归”模型，这里α是截距，而β是斜率。模型（1.2）关于X是高阶的且是抛物线方程。虽然Y与X间的关系是非线性的，我们仍然说模型（1.2）是线性的，这是由于参数α，β和γ是线性地出现的。模型（1.3）是多元线性回归模型中最简单的例子，它有两个回归变量X_1和X_2。

（1.1）—（1.3）的回归模型可以用数学记号写成种种形式。仅就模型（1.1）而言，下面各种可供选择的形式可能会遇到：

$$Y_t = \alpha + \beta X_t + \varepsilon_t \tag{1.4}$$

$$E(Y) = \alpha + \beta X \tag{1.5}$$

$$E(Y \mid X) = \alpha + \beta X \tag{1.6}$$

式（1.4）提醒使用者数据集由n个可用的响应值Y_t（$t = 1, 2, \cdots, n$）组成，它们在独立变量X_t（$t = 1, 2, \cdots, n$）的相应值处测出。这n对观察值可以写成(X_1, Y_1)，(X_2, Y_2)，$\cdots$，(X_n, Y_n)。对于每一对(X_t, Y_t)，存在一个“干扰”ε_t，它是由于Y_t偏离它的期望值$\alpha + \beta X_t$而引起，已知$E(\varepsilon) = 0$。（1.5）和（1.6）的左端表示隐含在使用模型（1.1）的一种看法，即对于一个指定的X值，Y的期望值是$\alpha + \beta X$。我们也可以简单地写这个模型为

$$Y = \alpha + \beta X \tag{1.7}$$

仅仅指出它的确定部分。

对本书中用到的所有模型，无论是线性的或非线性的，模型中未知参数的估计都是以最小二乘准则（LS）为依据的。有其它的准则可用，但在本著作的范围内将不探索它们。只要说明在适当的条件下，最小二乘准则有某些最优的性质也许就够了。这些最优的性质的详细讨论可参阅Malinvaud的书（1970,第3章和第5章）。在本书里，往后我们假定ε_t是独立同分布的正态随机变量（iidN），具有零均值和有限方差σ^2，在这种情况下，线性模型的未知参数的LS估计也是该模型参数的最大似然（ML）估计。它们是无偏的，且在线性正则无偏估计类中具有最小方差的性质。假定上面的要求是满足的，那末最小二乘准则提供了实践中最佳可用的估计。

让我们考察很简单的数据集（表1.1），它对说明线性与非线性的差异是有用的。当仅仅有两次观察，即$n=2$，数对（X_t，Y_t）是（2，2.5）和（3，10.0）。这些数据可用图1.1表示。这个图对于每一个t值，如同通常的Y_t对X_t的图一样，假定从得出这些数据的系统或过程的理解有理由相信可以使用“通过坐标原点的直线回归”的简单模型，

即

$$Y_t=\beta X_t+\varepsilon_t \tag{1.8}$$

β的LS估计是由观察到的Y_t与假定的真实模型的偏差平方的和取极小值而得到，即极小化

$$S(\beta)=\sum_{t=1}^{n}(Y_t-\beta X_t)^2 \tag{1.9}$$

表1.1　说明性的数据

观察数	X	Y
1	2	2.5
2	3	10.0

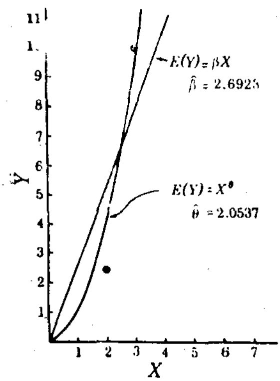


图1.1　表1.1的数据与模型（1.8）和（1.9）的拟合

以后，为了简洁起见，求和限将从Σ中省略。用S代替$S(\beta)$以简化记号。S的极小值可由（1.9）式对β求微分得到，让导数等于零，并解出β，它的解用$\hat{\beta}$表示，称它为β的LS估计。因此由

$$\frac{\partial S}{\partial \beta}=-2\sum(Y_t-\beta X_t)X_t=0$$

得到如下解：

$$\hat{\beta}=\frac{\Sigma X_t Y_t}{\Sigma X_t^2} \qquad (1.10)$$

当假定X_t没有误差并且它不是随机变量时，$\hat{\beta}$是随机变量Y_t的线性组合.如果Y_t假定是围绕其均值βX_t波动且具有有限（虽然未知）方差σ^2（即ε_t的方差）的正态分布，则由此得出$\hat{\beta}$也是正态分布.此外，$\hat{\beta}$的期望是β，这意味着$\hat{\beta}$是β的无偏估计，并且$\hat{\beta}$的方差是

$$\mathrm{Var}(\hat{\beta})=\frac{\sigma^2}{\Sigma X_t^2}$$

它是β的任何线性无偏估计中具有最小的可能方差．对于线