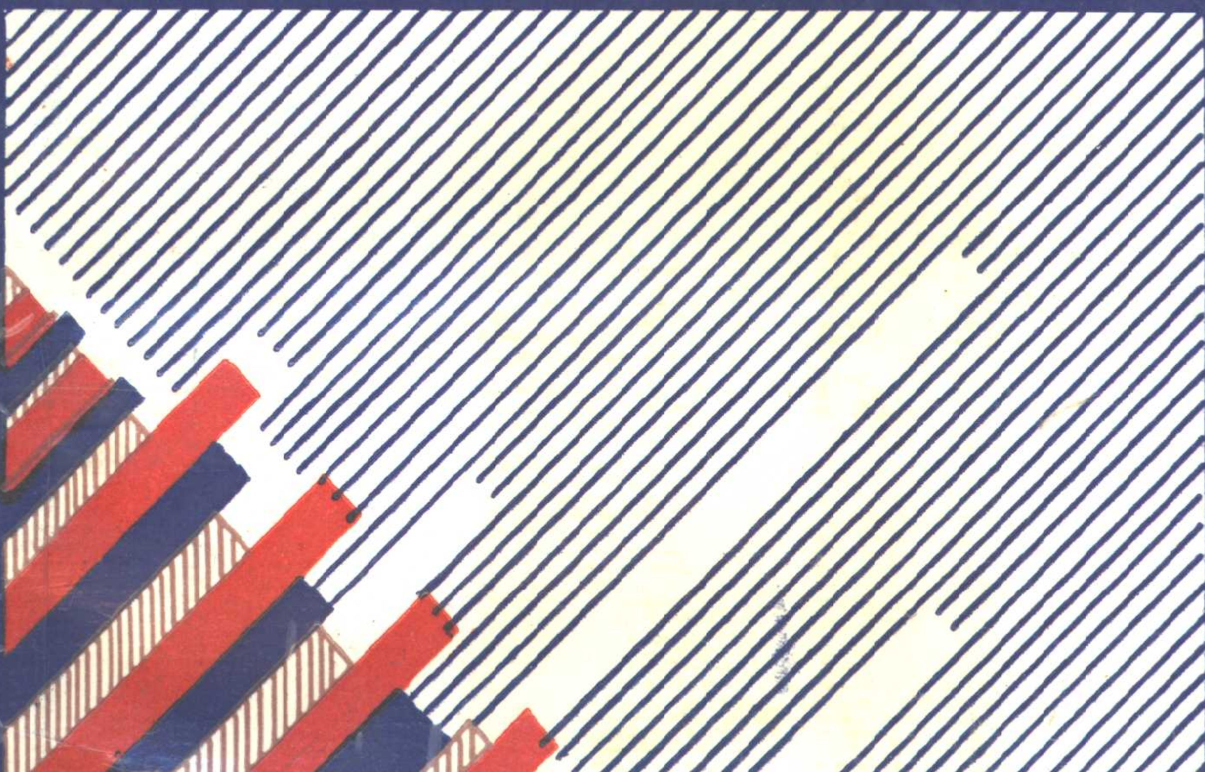
近代非线性回归分析

韦博成 著



东南大学出版社

近代非线性回归分析

韦博成 著

（国家自然科学基金资助项目）

东南大学出版社

1989·南 京

内 容 简 介

本书论述近代非线性回归分析的原理、方法和某些应用，侧重介绍以曲率度量为基础的几何理论，总结了近年来的若干新进展，内容包括常用非线性模型、最小二乘估计量、模型的曲率度量、置信域的曲率表示、估计量的渐近矩及其应用、带约束的非线性模型以及与曲率有关的渐近理论等。

本书叙述力求深入浅出，只要具有大学数理统计知识就能阅读本书内容。本书可作理工科大学生和研究生的教学参考书，亦可供有关教师、科技人员和统计工作者参考。

责任编辑　徐步政

近代非线性回归分析

韦博成　著

东南大学出版社出版

（南京四牌楼2号）

江苏省新华书店发行　东南大学印刷厂印刷

开本787×1092毫米 1/32　印张8 5/16字数 187千

1989年3月第1版　1989年3月第1次印刷

印数：1—3000册

ISBN 7-81023-125-1

O·23　　　定价：3.40元

序　　言

现实世界中，严格的线性模型并不多见，它们或多或少都带有某种程度的近似。随着科学技术和近代统计学的飞速发展，不能简单化为线性回归的非线性模型越来越多地出现在统计学家面前。农业，生物，经济，工程等各部门都提出许多非线性回归以及其它非线性统计问题。因此，积极开展非线性模型的研究在理论与实践中日趋重要。非线性模型是线性模型的自然推广，也是必然发展趋势。

非线性模型的研究始于六十年代初期，但近二十年中进展并不很快。直到1980年，加拿大统计学家 Bates 和 Watts 引入曲率度量以后，才得到较快的发展。本书着重介绍1980年以来在非线性回归模型的理论和方法上的新进展，并突出几何概念和几何方法的运用。

本书共分五章合三个附录。第一章概括介绍经典理论，列举常见的非线性模型，并介绍最小二乘估计量的计算方法和渐近性质。第二章到第五章详细介绍近年来非线性回归模型的新进展，系统总结了美、加统计学家Bates, Watts, Hamilton, Cook, Tsai 等人以及著者本人的研究成果。第二章引入非线性强度的曲率度量概念、计算公式以及对各种具体模型的应用。在本章中还第一次提出带有约束的非线性模型（并在第三章、第四章得到进一步发展）。第三章利用统计曲率研究置信域问题，其中包括带有多余参数和线性约束的情形。第四章提出利用随机展开式计算各种矩的统一方法以及对渐近理论、预测问题和诊断问题的应用。第五章结合非线性模型讨论由日、美统计学家 Amari, Efron 以及 Hinkley 等人发展起来的与

统计曲率有关的渐近理论。

本书内容较新，但对读者基础知识的要求并不高，只要了解大学本科数理统计的内容，就能阅读本书。为了便于读者阅读，本书还准备了三个附录，它们是了解正文所必需的。特别是附录1，介绍立体阵的运算和性质，在本书正文中经常被引用，读者必须十分熟悉。

本书初稿曾对本校研究生以及全国青年教师讲习班进行过几次讲授。讲授过程中曾不断进行修改和补充。不少同志提出了宝贵意见，华东师范大学魏宗舒教授，武汉大学张尧庭教授，清华大学肖树铁教授以及东南大学高金衡教授对本书的出版一直表示极大的支持与关心，在此一并表示衷心的感谢，并对东南大学出版社以及孙文治、徐步政同志的大力支持表示由衷的谢意。

由于作者水平所限，难免有不妥和谬误之处，恳请同行专家及广大读者不吝赐教。

著　　者

1988年4月

符 号 说 明

$\triangleq$	“定义为”或“记为”
Y	n 维观测向量
ε	n 维随机误差向量
θ	p 维参数向量
Θ	p 维参数空间
$f(\theta)$	$\triangleq f(X,\theta)$，模型函数
$e(\theta)$	$\triangleq$ n维向量$Y-f(\theta)$
$\|a\|$	向量 a 的欧氏模
$S(\theta)$	$\triangleq \|e(\theta)\|^2=\|Y-f(\theta)\|^2$
$V(\theta)$	$f(\theta)$的一阶导数，$n\times p$阶矩阵
$W(\theta)$	$f(\theta)$的二阶导数，$n\times p\times p$阶立体阵
iid	独立同分布
iidN	独立同分布且服从正态分布
$\hat{\theta}$	θ 的最小二乘估计量
$\hat{e}$	$\triangleq Y-f(\theta)$，残差向量
$\triangle f$	$\triangleq f(X,\theta)-f(X,\theta)$，拟合误差向量
s^2	$\triangleq (n-p)^{-1}S(\theta)$，方差 σ^2 的估计量
$\eta=f(\theta)$	解轨迹，记为 π
T_θ	解轨迹在 θ 处的切空间
LSE	最小二乘估计量
MLE	最大似然估计量
E_n	n 阶单位矩阵（也记为E）
A'	矩阵A的转置

A^{-1}	矩阵A的逆
A^{+}	矩阵A的加号逆
$\|A\|$	矩阵A的行列式
$\mathrm{tr}(A)$	矩阵A的迹
$\mathrm{Vec}(A)$	矩阵A的向量化运算
P_A	矩阵A产生的投影阵
P_T	切空间的投影阵
P_N	法空间的投影阵
$[A][Z]$	矩阵A与立体阵Z的方括号乘法（参见附录1）
$\mathrm{tr}[Z]$	$n\times p\times p$阶立体阵的迹（n维向量）
$\mathrm{Vec}[Z]$	立体阵Z的向量化运算（见附录1）
Q	其列向量为切空间的标准正交基
N	其列向量为法空间的标准正交基
R	p阶上三角矩阵 $V=QR$
L	$\triangleq R^{-1}$，p阶上三角矩阵
U	$\triangleq L'WL$，$n\times p\times p$阶立体阵
I	$\triangleq[N'][U]$，$(n-p)\times p\times p$阶固有曲率立体阵
P	$\triangleq[Q'][U]$，$p\times p\times p$阶参数效应曲率立体阵
B	$\triangleq[e'][U]$，有效残差曲率阵
$l(y,\theta)$	对数似然函数
$\dot{l}(y,\theta)$	$l(y,\theta)$关于θ的一阶导数
$\xrightarrow{L}$	依分布收敛
$\mathscr{L}(A)$	矩阵A的列产生的向量空间
$\oplus$	向量空间的直接和
δ_{kl}	$\triangleq k$，l相等时为1，否则为0
$\otimes$	矩阵的Kronecker乘积
$J(Y)$	Y的 Fisher 信息阵

$G(\theta)$　　观察信息阵

$E(X)$　　X的数学期望

$\mathrm{Var}(X)$　X的方差

$\mathrm{Cov}(X,Y)$　X，Y之间的协方差

$\mathrm{Bias}(\theta)$　θ 的偏差

τ　　$\triangleq Q'\varepsilon$　p维向量

λ　　$\triangleq N'\varepsilon$　$n-p$维向量

目　录

第一章　非线性模型及其最小二乘估计

线性模型的一般形式为

$$y=x_1\beta_1+\cdots+x_p\beta_p+\varepsilon$$

其中 $x_1,\cdots,x_p$ 为可观测的已知变量，y 为可观测的随机变量，ε 为不可观测的随机误差项，$\beta_1,\cdots,\beta_p$ 为未知参数。通常可以进行 n 次观察得

$$y_t=x_{t1}\beta_1+\cdots+x_{tp}\beta_p+\varepsilon_t \quad (t=1,\cdots,n)$$

我们可以通过观察值 $y_1,\cdots,y_n$ 对模型及其参数进行统计分析。在一般情况下可以考虑

$$y=f(x,\ \theta)+\varepsilon \tag{1.1}$$

其中 x 为可观测的已知变量，可以是向量；y 为可观察的随机变量；ε 为不可观测的随机误差项；$\theta=(\theta_1,\cdots,\theta_p)'$ 为未知参数；f 称为模型函数，它的函数形式已知，但含有未知参数θ。如果 f 是 θ 的线性函数，则(1.1)式化为前面叙述的线性模型，否则就称为非线性模型。对于非线性模型，(1.1)式的 n 个观察值为

$$y_t=f(x_t,\ \theta)+\varepsilon_t \quad (t=1,\cdots,n) \tag{1.2}$$

与线性模型一样，我们的问题就是通过观察值 $y_1,\cdots,y_n$对模型及其参数进行统计分析。

§1.1 常用非线性模型

以下我们系统地介绍几类常见的非线性模型，更具体的讨论及数值例题可参见[23]，[47]。

一、生长模型

生长模型是农业，生物，经济，化工等领域中应用极广泛的一类模型。例如，草原牧草的再生过程就是一个典型的生长模型。牧草的产量 y 随时间 x 而增长，但它又受环境、天气等许多随机因素的影响，因而有 $y=f(x, \theta)+\varepsilon$ 的形式。这例子还是一种所谓 s 形生长模型，即其生长速度开始时不断增加，而增加到某一个转折点就逐渐减小，生长曲线呈 s 形（见图1.1)。其它如经济增长，化学反应等也常有这种规律。常用的模型有

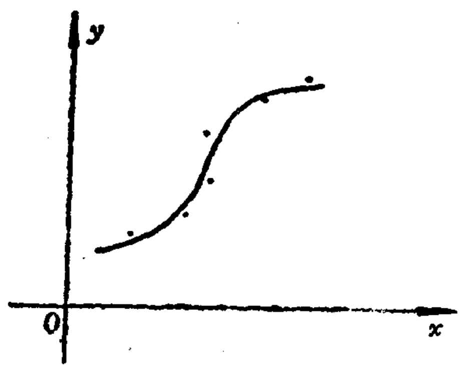


图1.1 s 形生长曲线

$$y=\exp(\theta_1-\theta_2\cdot\theta_3^x)+\varepsilon \qquad \text{(Gompertz模型)}$$

$$y=\frac{\theta_1}{1+\exp(\theta_2-\theta_3^x)}+\varepsilon \qquad \text{(Logistic模型)}$$

$$y=\theta_1-\theta_2\exp(-\theta_3x^{\theta_4})+\varepsilon \qquad \text{(Weibull模型)}$$

在生长模型中又可分为所谓经验型和机械型。经验型即根

据行之有效的经验公式得到。以上各式都可认为是经验型。机械型是根据某种规律或假定求出微分方程或差分方程的解，再分析考虑随机因素的影响而得到。我们来看几种常见的机械型生长模型。

设 y 的生长极限为 α，而生长速度与"生长余量" $\alpha-y$ 成正比。这时 y 满足微分方程

$$\frac{dy}{dt}=k(\alpha-y)$$

由此解出 y，得到模型

$$y=\alpha(1-\beta e^{-kt})+\varepsilon \quad \text{(单层增长型)} \qquad (1.3)$$

例如，桔子树干的周长 y 与生长天数 t 的关系就常采用这个模型。

假定 y 的相对增长速度与相对"生长余量"成正比，则 y 满足微分方程

$$\frac{dy}{dt}\Big/y=k(\alpha-y)/\alpha$$

由此解出 y，得到模型

$$y=\alpha(1+\beta e^{-kt})^{-1}+\varepsilon \quad \text{(对数型)}$$

此模型称为对数型的原因是，如果取乘法误差，即 $y=\alpha(1+\beta e^{-kt})^{-1}(1+\varepsilon)$，并取对数 $z=\log y$，则得到

$$z=\gamma-\log(1+\beta e^{-kt})+e \quad \text{(对数型)}$$

在对数模型中，虽然对随机误差项 ε 或 e 需要一些限制，但实用上则是经常采用的方法。

若假定相对速度与"对数余量" $\log\alpha-\log y$ 成正比，则 y 满足微分方程

$$\frac{dy}{dt}\Big/ y=k(\log \alpha-\log y)$$

解出 y 可得到模型

$$y=\alpha \exp(-\beta e^{-kt})+\varepsilon \quad \text{（Gompertz 模型）}$$

Gompertz 模型是一个 s 形增长模型，它在经济学中有广泛的应用[47]。如果令$z=\log y$，并取乘法误差，则模型化为

$$z=\gamma-\beta e^{-kt}+\varepsilon \quad \text{（渐近回归模型）} \qquad (1.4)$$

这不是 s 形增长模型，其形式与单层增长型(1.3)式类似。

二、产量密度模型

设 x 表示某种作物的种植密度，即单位面积的种植株数，y 表示每株作物的产量，则作物单位面积的产量就是xy。x 确定以后，则单位面积产量取决于 y，而 y 则受到天气、环境等多种因素的影响，一般可表示为 $y=f(x,\theta)+\varepsilon$，其中 θ 为待定参数，ε 为随机误差。模型函数 f 与作物的类型、地区等等有密切的关系。以下 Holliday 模型是效果比较好的常用模型

$$y=(\theta_1+\theta_2 x+\theta_3 x^2)^{-1}+\varepsilon \quad \text{（Holliday 模型）}$$

例如谷物、洋葱、胡萝卜等作物经常可采用这种模型。

当上式中取 $\theta_3=0$，则称为渐近模型，即

$$y=(\theta_1+\theta_2 x)^{-1}+\varepsilon$$

这也是常用的。

三、渐近回归模型

最简单的渐近回归模型为

$$y=\alpha-\beta\gamma^x+\varepsilon \quad \text{（渐近回归模型）}$$

其中 $0<\gamma<1$。这在生物及工程中都是常用的模型，如化学反应中反应物的分解量 y 与时间 x 的关系；鱼的长度 y 与其年龄 x 的关系等都经常采用这种模型。(1.3)、(1.4) 也是渐近回归模型。这些模型不具有 s 型生长规律，而生长极限将趋于一个稳定的值。其它常用的渐近回归模型有

$$y=\frac{1}{\alpha}-\beta\gamma^x+\varepsilon$$

$$y=\alpha-\exp\{-(\beta+\gamma^x)\}+\varepsilon$$

四、其它模型

在非线性模型的文献中，以微分方程的解，特别是线性微分方程组的解提出非线性模型问题是很常见的（可参见[10]，[31]，[34]）。下列模型是一个很典型的例子。

例1.1 一级不及逆化学反应。

假定在一个连续的化学反应中，物质 x 以速度 θ_1 分解为物质 y，而物质 y 又以速度 θ_2 分解为物质 z。今假定 θ_1 和 θ_2 为未知参数，又假定在 t 时刻物质 x,y，z 的含量分别为 $x(t)$，$y(t)$，$z(t)$。它们满足一级化学反应方程，即

$$\frac{dx}{dt}=-\theta_1 x$$

$$\frac{dy}{dt}=\theta_1 x-\theta_2 y$$

$$\frac{dz}{dt}=\theta_2 y$$

其初始条件可设为 $x(0)=1$，$y(0)=z(0)=0$。参数空间满足 $\theta_1\geqslant\theta_2$（即假定反应开始时仅有物质 x 为一个单位，且 x 分解

为 y 的速度比 y 分解为 z 的速度要快)。

由此很容易求得 $x(t)$，$y(t)$，$z(t)$的解。由于参数 θ_1，θ_2 未知，且有各种随机干扰。因此最后得到下列非线性模型（[10]曾经对此模型作过详细研究）

$$\begin{cases} x=e^{-\theta_1 t}+\varepsilon_1 \\ y=(\theta_1-\theta_2)^{-1}\theta_1(e^{-\theta_2 t}-e^{-\theta_1 t})+\varepsilon_2 \\ z=1-(\theta_1-\theta_2)^{-1}(\theta_1 e^{-\theta_2 t}-\theta_2 e^{-\theta_1 t})+\varepsilon_3 \end{cases} \quad (\text{当 } \theta_1>\theta_2)$$

以及

$$\begin{cases} x=e^{-\theta_1 t}+\varepsilon_1 \\ y=\theta_1 te^{-\theta_1 t}+\varepsilon_2 \\ z=1-e^{-\theta_1 t}-\theta_2 te^{-\theta_1 t}+\varepsilon_3 \end{cases} \quad (\text{当 } \theta_1=\theta_2)$$

由以上表达式可知，本例还是一个多元非线性模型。

在实用上，非线性模型的形式是多种多样的，为了说明应用的广泛性，我们再列举一些实例如下。

热敏电阻的传导系数 y（即电阻的倒数）与温度 x 的关系可表示为

$$y=\theta_1\exp\{\theta_2(x+\theta_3)^{-1}\}+\varepsilon$$

软铸铁的伸长率 x 与张应力 y 和屈应力 z 的关系可表示为

$$y=\theta_1+\theta_2 x^{-\theta_3}+\varepsilon_1$$

$$z=\theta_4+\theta_5 x^{-\theta_6}+\varepsilon_2$$

某种化学产品的反应速度 y 与催化剂的含量 x_1 及反应物的含量 x_2 之间的关系可表示为

$$y=\theta_1\theta_3 x_1(1+\theta_1 x_1+\theta_2 x_2)^{-1}+\varepsilon$$

饱和蒸汽的压力 y 与温度 t 的关系可表示为