

科学 专著丛书

非线性 时间序列分析

NON-LINEAR TIME
SERIES ANALYSIS

安鸿志 陈 敏 著

上海科学技术出版社

科学专著丛书

非线性时间序列分析

安鸿志 陈 敏 著

上海科学技术出版社

责任编辑 赵序明

科学专著丛书

非线性时间序列分析

安鸿志 陈 敏 著

上海科学技术出版社出版、发行

(上海瑞金二路 450 号)

新华书店上海发行所经销 常熟市印刷四厂印刷

开本 787×1092 小 1/16 印张 20.75 插页 4 字数 332 000

1998 年 8 月第 1 版 1998 年 8 月第 1 次印刷

印数: 1—1 200

ISBN 7-5323-4373-1/O · 207

定价: 38.00 元

本书如有缺页、错装或坏损等严重质量问题,
请向承印厂联系调换

内 容 提 要

本书概述了非线性时间序列分析的基本内容以及近代的重要进展。全书分为两部分:第一部分为模型概论,即第1章至第4章,其中包括了:比较线性与非线性时间序列;概述与非线性时间序列有关的非线性动态系统;介绍各种非线性时间序列模型;论述这些模型的平稳解、遍历性、高阶矩和可逆性。第二部分介绍统计方法,即第5章至第9章,包括非线性时间序列预报方法;非线性性检验方法;非线性时间序列建模方法;与其相关领域(如混沌、分形和神经网络)之间的联系;若干实例应用分析。

本书适合于作大学本科高年级学生和研究生的教材或参考书;也是从事时间序列分析研究工作人员及使用时间序列方法分析数据资料工作者的重要参考书。

《科学专著丛书》序

如果说科学研究论文是创造性科学工作的发表性记录,那末科学技术学术专著则是创造性科学工作的总结性记录。前者注重的是优先权,后者注重的是系统化。

在大量科学研究的基础上,对一个专题或一个领域的研究成果,作系统的整理总结,著书立说,乃是科学研究工作不可少的一个组成部分。著书立说,既是丰富人类知识宝库的需要,也是探索未知领域、开拓人类知识新疆界的需要。特别是在科学各门类的那些基本问题上,一部优秀的学术专著常常成为本学科或相关学科取得突破性进展的基石。所以,科学技术学术专著的著述和出版是一项十分重要的工作。

近20年来,中国的科学事业有了迅速的发展,涌现了许多优秀的科学研究成果,为出版学术专著提供了坚实的基础。值此20世纪90年代,在出版学术专著方面,中国的科学界和出版界都在抓紧为本世纪再加些积累,为迎接新世纪多作些开拓。我高兴地看到,作为这种努力的一个部分,《科学》杂志的出版者——上海科学技术出版社推出了这套《科学专著丛书》。

上海科学技术出版社是科学技术界熟悉和信赖的一家出版社,历来注重科学技术学术专著的出版。《科学》杂志的编者组织编辑学术系列丛书,也不是第一次。在本世纪三四十年代,就曾推出过《科学丛书》,其中不乏佳作,对当时的学术研究起了很好的作用。

《科学》在中国是一份历史最长的综合性科学刊物,80年来与科学技术界建立了广泛的密切联系。现在推出的这套《科学专著丛书》正是这种联系的产物。我相信,加强这种联系,著者与编者、出版者,科技界与出版界共同努力,精心选题,精心编辑,精心出版,一定能使这套专著丛书反映

出中国科学技术研究的最新水平,为本世纪多留下几本中国学者的优秀专著,为迈向新世纪多铺下几块引路的基石!

周光召

(《科学》杂志编委会主编)

1994年8月

前 言

在过去的半个多世纪里,时间序列分析得到了迅速发展。特别是对线性时间序列分析的研究,已经取得了系统和丰富的成果。但是对于非线性时间序列分析的研究,仅在近二十多年里才逐渐被重视起来。

线性时间序列仅是时间序列的一种特殊类型,而且是最简单的一类。尽管如此,五十多年来,被研究得最深入,应用也最广泛的,也仅是有限参数型线性时间序列模型,例如熟知的自回归滑动平均模型。所以说,非线性时间序列分析是一个极其广泛的研究范畴,有着极为丰富和深刻的研究内容。

非线性时间序列,哪怕是很简单的非线性时间序列,它们能显示出的奇异现象,足以令人新奇不已。况且,非线性时间序列分析与许多非线性科学领域有着密切的联系,例如与混沌、分形和神经网络等领域均有密切的关系。这说明了非线性时间序列能表现出更丰富、更复杂的客观现象,比线性时间序列有着更广阔的应用背景。实际上,在很多领域,例如在经济、环境、生物和气象等领域,已经显示出非线性时间序列分析的应用背景。因此,近年来对非线性时间序列分析的研究受到广泛关注。

研究非线性时间序列,很难仿效或沿用已有的线性时间序列的理论与方法,而且所面临的困难也远远超过线性情况。正像在其他非线性科学领域一样,人们遇到了许多新的激动人心的挑战。

近年来,关于非线性时间序列分析的文献,正在不断丰富。本书作者在了解有关文献的同时,也从事了某些理论和应用研究。并且在非线性时间序列模型的平稳解、遍历性等理论方面,以及非线性性检验方法和随机条件方差模型的应用方面,取得了某些研究成果。这些研究工作使作者对非线性时间序列分析的丰富内涵,有了较新的和深入的了解。这促使作者写出此书,以期吸引更多的同仁学者关心非线性时间序列分析的应用与发展。

在此,我们要感谢汤家豪教授的鼓励和帮助,提供许多最新的论著的

预印本,使我们较及时地了解非线性时间序列研究的最新进展。作者还要感谢香港中文大学统计系王永雄教授,受他的邀请,本书第一作者曾在该系研究生的讨论班上讲过本书的多数章节,并在该系访问期间完成本书初稿。作者还要感谢曾与他们共同组织参加非线性时间序列讨论班的同事们,书中许多内容曾与他们交流或讨论过。此外,还应当感谢北京大学概率统计系谢衷洁教授,他细心审阅了书稿,并提出许多宝贵意见。

该书第一作者多年受到国家自然科学基金委的资助,此书是属于天元重点基金资助项目的科研成果之一,一并在此致谢。

由于作者知识所限,书中谬误之处难免,恳望学者同仁不吝批评指正。

作 者

1996年3月于北京中国科学院

应用数学研究所

符号说明

γ_k	宽平稳时间序列的自协方差
$\gamma_{k,j}$	时间序列的三阶矩函数
$\hat{x}_n(l)$	l 步条件期望预报
$\hat{x}_n^*(l)$	l 步投影预报
$\ \cdot\ $	Euclid 范数
$\ \cdot\ _v$	与 Euclid 范数 $\ \cdot\ $ 等价的某一向量范数
$T_t(x_0)$	以 x_0 为初始值用变换 T 迭代 t 步的值
$L(x_0, T)$	迭代序列 $\{T_t(x_0), t \geq 1\}$ 的所有可能的极限值全体
$L(T)$	变换 T 的吸引集
$r_t(x_0, B)$	点 $T_t(x_0)$ 到集合 B 的普通 Euclid 距离
$A_T(B)$	变换 T 的吸引元 B 的吸引域
$A \geq B$	$A - B$ 是半正定 Hermite 矩阵
$\rho(A)$	矩阵 A 的谱半径, 即 A 的特征值最大模
A^T	矩阵 A 的转置
R^m	m 维实数域空间
$\mathcal{B}_m$	R^m 中的 Borel σ -代数
$P(x, A)$	$(R^m, \mathcal{B}_m)$ 上的转移概率
$I(x \in A)$	集合 $A \in R^m$ 的示性函数
$\ F - G\ _r$	分布 F 与 G 之差的全变差
$\text{Vec}(A)$	矩阵 A 的拉直运算
$A \otimes B$	矩阵 A 与 B 的 Kronecker 积
$\xrightarrow{p}$	依概率收敛
$\xrightarrow{d}$	依分布收敛
a. s.	概率为 1 地成立

$d_k(A)$	R^m 中有界集合 A 的 Kolmogorov 容度
$\triangleq$	表示“定义”或“记为”
$[x]$	取实数 x 的整数部分
$\mathcal{F}_{a,b}^x$	由 $\{x_s, a \leq s \leq b\}$ 生成的 σ -代数
$\mathcal{F}_t^x$	由 $\{x_s, s \leq t\}$ 生成的 σ -代数
$\epsilon_t \mathcal{F}_{t-1}^x$	给定 $\mathcal{F}_{t-1}^x$ 的条件下, ϵ_t 的条件分布
μ_m	R^m 空间的 Lebesgue 测度

目 录

《科学专著丛书》序

前 言

符号说明

第 1 章	浅释线性与非线性时间序列.....	1
§ 1.1	平稳时间序列	1
§ 1.2	线性时间序列	4
§ 1.3	非线性时间序列	6
§ 1.4	平稳序列的预报及新息序列的假定	9
§ 1.5	线性与非线性自回归模型.....	12
第 2 章	离散动态系统概述	17
§ 2.1	线性离散动态系统.....	17
§ 2.2	动态系统的吸引元和吸引域.....	22
§ 2.3	动态系统的稳定性.....	36
第 3 章	非线性时间序列模型	59
§ 3.1	时间序列分类概述.....	59
§ 3.2	可加噪声模型.....	62
§ 3.3	随机条件方差模型.....	72
§ 3.4	非线性自回归滑动平均模型.....	79
§ 3.5	其他非线性时间序列模型.....	81
第 4 章	非线性时间序列模型的概率结构	85
§ 4.1	一般状态马氏链.....	85
§ 4.2	非线性自回归模型的几何遍历性.....	95
§ 4.3	几何遍历的 NLAR 模型的例子	117

§ 4.4	非线性时间序列模型的平稳解	125
§ 4.5	非线性时间序列模型的高阶矩	150
§ 4.6	时间序列的两种可逆性	157
第 5 章	非线性时间序列的预报	165
§ 5.1	时间序列预报的概述	165
§ 5.2	可加噪声 NLAR 模型的预报	175
§ 5.3	随机条件方差模型的预报	180
§ 5.4	非线性 ARMA 模型的预报	185
第 6 章	时间序列的线性性检验	190
§ 6.1	检验的分类	190
§ 6.2	Kolmogorov-Smirnov 检验方法	193
§ 6.3	混合型检验方法	199
§ 6.4	其他非参数型检验方法	203
§ 6.5	Lagrange 乘子检验方法	206
第 7 章	模型的估计	214
§ 7.1	参数型非线性模型的估计	214
§ 7.2	非参数估计	243
§ 7.3	模型阶数的估计	250
第 8 章	非线性时间序列与相关领域的联系	254
§ 8.1	非线性时间序列与混沌	254
§ 8.2	经验分布与伪随机序列	266
§ 8.3	非线性时间序列与分形维数	273
§ 8.4	时间序列与神经网络	281
第 9 章	实例分析	287
§ 9.1	加拿大山猫数据的统计分析	287
§ 9.2	太阳黑子数据的统计分析	295
§ 9.3	其他实例分析	301

参考文献.....	306
名词索引.....	313

第 1 章

浅释线性与非线性时间序列

本章的目的是简要地介绍线性时间序列与非线性时间序列的联系,彼此相似之处以及两者间的实质性差异.通过对线性时间序列基本描述方法的简略回顾以及与非线性时间序列的比较,概述研究非线性时间序列所涉及的基本内容,现在所使用的一些描述方法以及所面临的主要困难.

§ 1.1 平稳时间序列

从较广的意义上说,时间序列是指被观测到的依时间次序排列的数据序列.从概率论角度来看,用随机过程来描述最为合适.随机过程被定义为一簇随机变量,即 $\{x_t, t \in T\}$,其中 T 表示时间 t 的变化范围.对每个固定的时刻 t , x_t 是一个一元随机变量,这些随机变量的全体就构成一个随机过程.当 $T = \{0, \pm 1, \pm 2, \dots\}$ 时,随机过程 $\{x_t, t \in T\}$ 可写成 $\{x_t, t = 0, \pm 1, \pm 2, \dots\}$,称之为随机序列.由于 t 代表时间,所以此类随机序列也称为时间序列.这种定义对线性与非线性时间序列都是适用的.

由上述可见,时间序列是一类特殊的随机过程——离散时间的随机过程.因此,时间序列的概率结构是被其有限维分布簇所确定的.粗略地说,当对每个自然数 m 和每组整数 $t_1, t_2, \dots, t_m, (x_{t_1}, x_{t_2}, \dots, x_{t_m})$ 的 m 维联合分布

$$P\{x_{t_1} < z_1, x_{t_2} < z_2, \dots, x_{t_m} < z_m\} = F(z_1, \dots, z_m; t_1, \dots, t_m)$$

都被给定时,时间序列 $\{x_t\}$ 的概率结构就被完全确定.详细的严格定义见文献[125].这一定义对线性与非线性时间序列都适用.

平稳时间序列,又简称平稳序列,是一类最重要的时间序列.对这类时间序列,其统计分析方法已有相当丰富的知识,应用也最为广泛.而且,许多非平稳序列的统计分析方法也是借助于与平稳序列的联系而建立起来的.例如,平稳序列的求和序列为非平稳序列,对它的研究即可借用平

稳序列的分析方法. 所以, 平稳序列在时间序列分析中起着重要作用. 时间序列的平稳性概念, 并不与时间序列的线性性直接联系.

定义 1.1.1 时间序列 $\{x_t\}$ 称为严平稳的, 如果对任何正整数 m 和整数 $t_1 < t_2 < \cdots < t_m$, 此序列中的随机变量 $x_{t_1+s}, x_{t_2+s}, \cdots, x_{t_m+s}$ 的联合分布与整数 s 无关, 即

$$F(z_1, \cdots, z_m; t_1 + s, \cdots, t_m + s) = F(z_1, \cdots, z_m; t_1, \cdots, t_m). \quad (1.1.1)$$

以上定义表明, 严平稳序列的任何有限维联合分布对于时间的平移是不变的. 关于时间序列的平稳性还有另一种不同意义的定义.

定义 1.1.2 时间序列 $\{x_t\}$ 称为宽平稳的, 如果 x_t 具有有穷的二阶矩, 而且满足以下条件:

- (1) $\mu_t = E x_t = c$;
- (2) $\gamma(t, s) = E(x_t - c)(x_s - c)$
 $= \gamma(t - s, 0).$

定义 1.1.2 表明, 宽平稳序列的均值与自协方差函数随时间的平移而不变. 另外, 其中条件(2)表明, 宽平稳序列的自协方差函数 $\gamma(t, s)$ 只与 $t - s$ 有关, 故 $\gamma(s, t)$ 实际上是被一元函数所确定. 以后, 我们用如下记号表示宽平稳序列的自协方差函数

$$\gamma_{t-s} = E(x_t - c)(x_s - c) = \gamma(t - s, 0).$$

由上述两个定义可知, 一个严平稳序列未必有二阶矩, 因而不一定是宽平稳序列. 反之, 一个宽平稳序列的分布未必随时间的推移而不变, 因此, 也不一定是严平稳序列. 为说明两种不同意义的平稳性的联系, 先介绍如下的正态时间序列的定义.

定义 1.1.3 如果时间序列 $\{x_t\}$ 的任何有限维分布都是多元正态分布, 则称 $\{x_t\}$ 为正态时间序列, 亦简称正态序列.

正态序列与平稳序列是两种不同的概念. 正态序列可以是平稳的, 也可以是非平稳的. 而且, 无论在严平稳或宽平稳意义下, 都是如此. 以下性质说明了严平稳与宽平稳的一般联系.

性质 1.1.4 如果 $\{x_t\}$ 是严平稳序列, 且 $E x_t^2 < \infty$, 则 $\{x_t\}$ 必是宽平稳序列. 反之, 不一定成立. 如果 $\{x_t\}$ 是正态序列, 则 $\{x_t\}$ 为严平稳序列与宽平稳序列是等价的.

此性质在时间序列分析中是熟知的结果. 关于宽平稳序列的自协方差函数的性质, 有如下熟知的结果.

性质 1.1.5 宽平稳序列 $\{x_t\}$ 的自协方差函数 γ_k 具有以下性质:

- (1) $|\gamma_k| \leq \gamma_0$;
- (2) $\gamma_k = \gamma_{-k}$;
- (3) 对任何自然数 m 和整数 $t_1 < t_2 < \cdots < t_m$ 及实数 $c_1, c_2, \cdots, c_m$, 有

$$\sum_{i=1}^m \sum_{j=1}^m c_i c_j \gamma_{t_i - t_j} \geq 0.$$

性质 1.1.6 $\{\gamma_k, k=0, \pm 1, \pm 2, \cdots\}$ 为某一宽平稳序列 $\{x_t\}$ 的自协方差函数, 当且仅当

$$\gamma_k = \int_{-\pi}^{\pi} e^{ik\lambda} dF(\lambda), \quad (1.1.2)$$

其中 $F(\lambda)$ 为一单调非降左连续函数, 且总变差是有穷的. 称 $F(\lambda)$ 为 $\{x_t\}$ 的谱分布函数, 而 (1.1.2) 式称为自协方差函数的谱表示.

定义 1.1.7 设 $\{x_t\}$ 是宽平稳序列, $F(\lambda)$ 是 $\{x_t\}$ 的谱分布函数, 如果存在非负可积函数 $f(\lambda)$, 使得

$$F(\lambda) = \int_{-\pi}^{\lambda} f(\lambda) d\lambda,$$

则称 $f(\lambda)$ 为 $\{x_t\}$ 的谱密度函数.

性质 1.1.8 设 $\{x_t\}$ 是宽平稳序列, 其自协方差函数 $\gamma(k)$ 满足

$$\sum_{k=-\infty}^{\infty} |\gamma_k| < \infty, \quad (1.1.3)$$

则 $\{x_t\}$ 必有非负谱密度函数 $f(\lambda)$, 且 $f(\lambda)$ 和 $\{\gamma_k\}$ 之间是 Fourier 变换的关系:

$$\begin{cases} f(\lambda) = \frac{dF(\lambda)}{d\lambda} = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{ik\lambda}; \\ \gamma_k = \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda. \end{cases} \quad (1.1.4)$$

以上性质的证明可见文献[20].

由上述性质可知, 宽平稳序列的自协方差函数与谱分布函数具有等价的描述能力, 它们都被序列的一维和二维分布 $F(z; t)$ 和 $F(z_1, z_2; t_1, t_2)$ 所决定. 反之, 除在正态时间序列情况下, 从自协方差函数不能唯一地决定序列的高维分布, 从而也就不能确定序列的全部概率结构. 所以说, 自协方差函数或谱分布函数都不是序列的全部概率结构信息的表现. 确切地说, 它们只是序列的一、二阶矩的结构表征.

一个最简单的例子是

$$x_t = \varepsilon_t \varepsilon_{t-1}, \quad (1.1.5)$$

其中 $\{\varepsilon_t\}$ 为独立正态 $N(0,1)$ 随机序列. 由(1.1.5)式易见,

$$Ex_t = E\varepsilon_t \varepsilon_{t-1} = E\varepsilon_t E\varepsilon_{t-1} = 0,$$

$$Ex_t^2 = E\varepsilon_t^2 \varepsilon_{t-1}^2 = E\varepsilon_t^2 E\varepsilon_{t-1}^2 = 1,$$

$$\begin{aligned} Ex_t x_{t+k} &= E\varepsilon_t \varepsilon_{t-1} \varepsilon_{t+k} \varepsilon_{t+k-1} \\ &= E\varepsilon_{t+k} E\varepsilon_t \varepsilon_{t-1} \varepsilon_{t+k-1} = 0 \quad (k \geq 1). \end{aligned}$$

以上诸式表明, (1.1.5)式的 $\{x_t\}$ 为宽平稳序列. 而依 $\{\varepsilon_t\}$ 的定义, 它也满足相同的性质, 即

$$E\varepsilon_t = 0, \quad E\varepsilon_t^2 = 1, \quad E\varepsilon_t \varepsilon_{t+k} = 0 \quad (k \geq 1),$$

即 $\{x_t\}$ 和 $\{\varepsilon_t\}$ 具有相同的均值和自协方差函数. 显然, 它们的分布结构不相同.

由于本书主要讨论严平稳性, 为了简便, 我们限定平稳性专指严平稳性. 本节所讨论的时间序列的平稳性概念, 无论对线性时间序列而言, 还是对非线性时间序列而言, 都是适用的, 并且具有重要意义. 对线性时间序列来说, 宽平稳序列起着重要作用, 而对非线性时间序列来说, 我们将更侧重于平稳序列的讨论.

§ 1.2 线性时间序列

相互独立且具有相同分布的随机变量序列是一类非常特殊的时间序列. 它在时间序列分析中占有非常重要的地位. 对此类序列还常假定为零均值和有穷方差, 在本书中我们称之为白噪声序列, 常用 $\{\varepsilon_t\}$ 表示之. 白噪声序列是一种概率结构十分简单的平稳序列. 许多重要的、有用的时间序列是由此类序列变换而来的. 本节将要讨论的线性序列就与白噪声序列有密切关系.

定义 1.2.1 如果随机序列 $\{x_t\}$ 可表为

$$x_t = \sum_{j=0}^{\infty} \beta_j \varepsilon_{t-j}, \quad (1.2.1)$$

其中系数序列 $\{\beta_j\}$ 满足

$$\sum_{j=0}^{\infty} \beta_j^2 < \infty, \quad (1.2.2)$$

而 $\{\varepsilon_t\}$ 是白噪声序列, 满足

$$E\varepsilon_t = 0, \quad E\varepsilon_t^2 = \sigma^2 < \infty, \quad (1.2.3)$$