时间序列分析基础

李先孝

华中理工大学出版社

时间序列分析基础

李 先 孝

华中理工大学出版社

时间序列分析基础
李 先 孝
责任编辑 龙 纯 曼
*
华中理工大学出版社出版发行
（武昌喻家山）
新华书店湖北发行所经销
华中理工大学出版社沔阳印刷厂印刷
*
开本：850×1168 1/32印张：8.25 字数：198 000
1991年11月第1版 1991年11月第1次印刷
印数：1—1 000
ISBN 7-5609-0609-5/O·86
定价：2.18元
（鄂）新登字第10号

内 容 简 介

本书是时间序列分析的基础读物，比较通俗系统地介绍了时间序列分析的理论、方法、应用及最新进展。本书共分八章，第一章是预备知识，介绍了概率空间、数理统计与随机过程的基础知识；第二章介绍了时间序列模型ARMA的类别及其特性；第三、四、五章介绍了时间序列分析方法及应用实例；第六章介绍了时间序列频域分析方法；第七、八章介绍了回归、自回归混合模型及离散动态系统辨识。书末附有参考习题，可供读者练习。

本书可作为工科院校有关专业师生及研究生学习《时间序列分析》的选修教材或参考书，也可供科技人员从事动态数据处理、分析、建模、预报和控制工作时参考。

序

随着计算技术的迅速发展，软硬环境的不断改善，时间序列分析方法应用范围日益广泛，倍受各个领域实际工作者的注意。所谓时间序列（以下简称时序），狭义地讲，是指按时间先后顺序排列的随机序列，或者说是定义在概率空间（Ω，F，P）上的一串有序随机变量集合，常记作

$$X_t,(t=0,\pm1,\pm2,\cdots);\tag{1}$$

它的每一个样本（现实）序列，是指按时间先后顺序对X_t所反映的具体随机现象或系统进行观测或试验所得到的一串有序动态数据，常记作

$$x_t\ (t=0,\pm1,\cdots).\tag{2}$$

通常为方便计，把X_t与x_t统记作x_t。时序x_t中对应不同时刻的随机变量之间是存在着某种统计联系的，这种联系使人们有可能根据随机变量“现在”和“过去”的观察值来推断随机变量“将来”的取值。所谓时序分析，简单地说，就是根据有序随机变量或者根据观测得到的有序数据之间相互依赖所包含的信息，用概率统计方法定量地建立一个合适的数学模型，并根据这个模型对相应的时序所反映的过程或系统作出预报或进行控制。当今，时序分析方法不断丰富，应用领域不断扩大，时序分析已经成为一门事实上相对独立的学科。

众所周知，严平稳序列未必存在二阶矩，因而它未必是宽平稳序列，反之亦然。但是，对于正态序列而言，它们却是一致的，而正态序列的分布唯一地由它的均值和方差确定。正因为如此，在时序分析中，是把平稳正态序列作为深入研究平稳序列的出发点的。起初，人们以时序$x_t(t=0,\pm1,\cdots)$的自协方差函数r_K或自相关函数ρ_K及其功率谱$s_K(f)$来反映它的统计特性。这就是

所谓的“非参数”模型．而近几十年发展起来的“参数化”模型，使时序分析的研究有了重要的突破．1927年首先由G.U.Yule提出了时序$x_t(t=0,\pm1,\cdots)$的AR模型[1]．AR模型是一种“参数化”模型，它是依据时序x_t构造的带有参数的随机差分方程，即

$$x_t-\varphi_1x_{t-1}-\cdots-\varphi_px_{t-p}=a_t, \tag{3}$$

式中，$\varphi_j(j=1,2,\cdots,p)$称为**参数**（实数），$a_t$是**白噪声**．这种模型一方面能够反映产生时序的动态系统的统计规律性，另一方面又能把时序的相互依赖性转化成相互独立的白噪声序列a_t．在这个基础上，逐步发展出了MA、ARMA等参数模型．对于零均值平稳正态序列(1)或(2)，作者较详细地介绍了Box-Jenkins方法[10]．

对于齐次非平稳时序$x_t(t=0,\pm1,\cdots)$，作者介绍了用差分或季节性差分等方法化非平稳时序为平稳时序过程．相应的模型有ARIMA(p,d,q)、ARIMA$(p,d,q)\times(P,D,Q)_T$等．

模型识别与参数估计是建模过程中的重要步骤．模型识别是根据(1)的一段样本，利用所谓相关函数ρ_K及偏相关函数φ_{KK}的截、拖尾特性及FPE、AIC、BIC等准则判定模型的类别和阶次，从而得到一个与(1)所反映的过程或系统尽可能相符的模型结构．然而，由于(1)的长度N毕竟是有限的，这就使我们面临两个问题：一是如何估计ρ_K、φ_{KK}？一是怎样才认为ρ_K、φ_{KK}是截、拖尾的？作者通过实例介绍了解决问题的方法．

在模型类别及阶次确定之后，我们自然要对未知参数φ_K、θ_K及σ_a^2进行估计，作者较详细地介绍了参数的矩估计、最小二乘估计以及AR模型参数的递推算法，其中包括Levinson递推算法[9]．

如果实际问题对于模型的精度要求不是很高，一般地讲，是可以把已通过模型识别、参数估计所得到的模型（称之为估计模型）用于预报或控制的．但是，严格地讲，还要从数学上对模

型进行适用性检验（也称为模型考核），即根据一定准则淘汰不"适用"的模型。作者介绍了除确定模型阶数的FPE、AIC、BIC准则外，还介绍了用于定阶的F-准则及白噪声独立性检验准则，同时，给出了模型检验准则的应用实例。

值得一提的是，作者用几个贯穿全书的实例，完整地描述了时序分析方法的主要步骤，这对于搞实际工作的读者来说是很有益处的。

预报是时序分析方法的重要应用之一，作者介绍了平稳线性最小方差预报以及各类模型的预报公式。

谱估计理论是时序分析方法的一个重要的侧面。实际上，用时域方法建立的三类模型，也可以从频域角度用有理谱来刻划。ARMA模型谱是与传统的Fourier谱不同的非传统谱，非传统谱理论的产生是时序分析方法发展过程中的一个重要突破[25]。1967年J.P.Burg提出了最大熵谱算法[2]。1968年E.Parzen提出了AR模型的谱估计[3]。1971年A.Van den Bos证明了AR模型谱与最大熵谱的等效性[4]。由于时序模型的谱仅依赖于有限个数参数φ_j、θ_j及白噪声a_t的方差σ_a，因此，谱估计的问题就转化为对模型参数的估计。只要模型类别、阶次、参数估计都是正确的，并且a_t是白噪声，那末据此计算的谱，就可以克服传统谱存在的能量泄漏、频谱分辨不清以及谱线偏移等固有缺陷。但随之而来的便产生了计算速度及精度问题，这就促使人们不断寻求保证精度的快速算法。J.P.Burg提出的最大熵谱估计理论，在这方面有很大改善，并在一定程度上克服了传统谱的弊病，作者对此做了较详细的介绍。同时，对S.L.Marple于1980年提出的算法[5]，以及我国的林代茂、茅于海于1983年提出的LUD算法也做了介绍[6]。为了比较传统谱的FFT算法与最大熵谱的Marple算法，作者采用了用FORTRAN语言编写的程序，根据仿真AR模型产生100个数据，用Marple算法程序进行了参数估计，结果表明，用Marple算法得到的谱较用FFT算法得到的谱优越（见本

书§6.2）。

应当指出，1978年美国威斯康星大学的吴贤铭和S.M.Pandit 提出一种较为切合工程实际的 DDS 算法[8]。近期又有人把时序分析与控制论中的状态空间法相结合，沟通了时序分析和现代控制理论之间的联系。

现代控制理论的一个重要组成部分是系统辨识。对于一个大型系统或者一个复杂的系统而言，虽然人们常常不能完全了解，甚至根本不可能了解它的结构，但是，其输入信息u_t与输出信息x_t常常是可以测量的。然而，对于一个连续系统，这是很困难的。计算技术的发展，使得用离散化的方法刻划和识别连续系统成为可能，于是解决离散系统的辨识问题日益引起人们的注意。作者主要讨论了单变元时不变线性离散动态系统的各种表达形式，即时序受控的自回归滑动平均模型CARMA（或ARMAX）及其特殊形式的辨识问题。

作者对近几十年来时序分析方法的重要成果作了不同程度的介绍，因而使本书有了一定的学术价值。另外，本书并不过多地追求理论上的严谨，而是更为注重介绍算法及应用实例，这对于工科专业的研究生以及实际工作者来说，不失为一本较好的读物。

王承中　于长春

前　言

时间序列分析是概率论、数理统计的一个分支。近十多年来，无论在理论、方法、或应用方面都有很大进展。本书较系统地介绍了一维时间序列分析的基础理论和方法，为了避免过多的涉及较深的数学理论，对有些定理、公式未予证明，只在书后列出了文献目录，供读者参考。读者只须具备（工科）《工程数学》基础知识，就可读懂全书。

本书是本人在吉林工学院研究生班开设《时间序列分析》课讲稿基础上修改补充而成的。在编写过程中，得到数学教研室大力支持，特别是王承中同志在百忙中阅读了初稿，并提出了许多中肯和建设性的修改意见，完稿后又认真细致地主审了全书并写了序；韩敬礼同志阅读了第六章和第八章并提出了宝贵意见，还提供了数值仿真例子和Marple程序；解治明同志热心地绘制了全部插图；王晓兵同学协助编制程序并计算了全部应用实例。在此，对他们表示诚挚的谢意。

由于本人水平所限，书中错误或欠妥之处在所难免，恳请读者批评指正。

李先孝　于长春

目　录

第一章 预备知识

假定读者已具备线性代数、概率论等基本知识．本章主要介绍工程数学没有涉及到的内容，如概率空间、数理统计与随机过程等，以便使其与时序分析内容更好地衔接起来．

§1·1 概率空间

由于时序$x_t(t=0,\pm1,\cdots)$是有序的随机变量集合，要分析它的特性，就必须研究随机变量的概率分布．为此，先从概率空间谈起．

所谓概率空间，是指公理化结构中三个基本概念Ω、F、P的总体，记作$(\Omega,\ F,\ P)$．其中Ω是样本空间，F是事件域，P是概率，并认为它们是预先给定的，至于怎样给定，要根据具体问题特点而定．首先给出样本空间Ω，指出所讨论的事件范围——事件域F，然后确定概率P的形式和数值．下面结合具体实例，构造概率空间．

1. 样本空间

在研究随机现象的随机试验（假定在相同条件下可以重复进行）中，每一个可能出现的结果，称为一个**样本点**（或**基本事件**）记作ω．全体样本点的集合，称为**样本空间**，记作$\Omega=\{\omega\}$．

例1 一枚硬币重复掷三次的试验，写出样本空间Ω．

用“0”表示出现正面，“1”表示出现反面，则可能结果是：

000；001；010；100；101；011；110；111．

分别记为ω_1；ω_2；ω_3；ω_4；ω_5；ω_6；ω_7；ω_8。
于是样本空间Ω是有限集合，

$$\Omega=\{\omega_i \mid i=1,2,\cdots,8\}.$$

例2 工人用机床加工某零件的试验。规定出现废品就停止加工，从加工开始至停止作为一次试验，构造该试验的样本空间Ω。

用"1"表示出现正品，"0"表示出现废品，则样本空间

$$\Omega=\{0,\ 10,\ 110,\ 1110,\ \cdots\},$$

或

$$\Omega=\{\omega_1,\ \omega_2,\ \omega_3,\ \cdots\}.$$

Ω是有限集，或可列无限集。

例3 某人作打靶的试验。规定打靶成绩用命中点到靶心距离x表示。那么一切可能中靶的结果构成区间$(0,\ r)$ （$r>0$常数，表示圆形靶半径），则样本空间

$$\Omega=\{x \mid x\in(0,r)\},$$

是不可列集。

2. 事件、事件域

(1) 事件

随机试验中的事件，是指样本空间Ω的子集合。如在例1（重复掷硬币三次的试验）中，以A表示出现两次正面一次反面的事件，即$A=\{001,010,100\}$；以B表示第一次正面，第二、三次反面的事件，即$B=\{011\}$；以$\bar{A}$表A不出现的事件，称为A的**余事件**，即$\bar{A}=\{000,101,011,110,111\}=\Omega-A$，
显然，事件A、B、$\bar{A}$都是样本空间Ω的子集合。

又如在例3（打靶试验）中，以A表示命中点离中心距离小于3的事件；以B表示命中点离中心距离大于2小于等于8的事件；即

$$A=\{x \mid x\in(0,3)\},$$

$$B=\{x\,|\,2<x\leqslant 8\}.$$

事件A、B都是样本空间Ω的子集.

(2)　事件域

定义　由样本空间Ω的某些子集所组成的集合F，如果满足下列条件：

1°　$\Omega\in F$；

2°　若$A\in F$，则$\overline{A}\in F$；

3°　若$A_n\in F(n=1,2,\cdots)$则和运算

$$\bigcup_{n=1}^{\infty}A_n\in F.$$

则称F是Ω上的一个**σ-域(σ-代数)**或**事件域**. F中的元素称为**事件**. 样本空间Ω(也是事件)称为**必然事件**. $\overline{\Omega}=\varnothing$ ($\varnothing$表示空集) 称为**不可能事件**. 显然，$\varnothing\in F$.

注意：事件域的定义只要求对求余、求和是封闭的，可以证明，事件域对求差、求积也是封闭的.

这里，事件域F是由Ω的某些子集构成的集合. 一般地，对由有限个或可列个样本点组成的样本空间Ω，则取Ω的全体子集构成的集合作为事件域F. 例如在例1中，由$\Omega=\{\omega_i\,|\,i=1,2,\cdots,8\}$的全体子集组成事件域$F$ (共有2^8个元素)，其中A、B、$\overline{A}$都是F的元素.

若Ω是一维区间(有限或无限)，则通常取F为包含Ω的一切子区间的一个最小的σ-域. 这个最小的σ-域叫做**Ω上的一维波雷耳(Borel)域**，其中的集合称为**波雷耳集**. 仿此，可定义n维波雷耳域.

3. 概率

定义　设Ω是给定的样本空间，F是Ω上的事件域(σ-域). 如果存在F上的一个集函数$P(A)$ $(A\in F)$满足下列条件：

1°　对每一个$A\in F$，$0\leqslant P(A)\leqslant 1$；

2° $P(\Omega)=1$;

3° 若$A_i\in F(i=1,2,\cdots)$且两两互不相容，即

$$A_i\cap A_j=\varnothing\,(i\neq j),$$

有

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i) \tag{1·1·1}$$

成立，则称$P(A)$是F上的一个概率测度，简称概率。

如例1(一硬币重复掷三次)中，已知样本空间Ω由8个样本点组成，即

$$\Omega=\{000,001,010,100,101,011,110,111\}.$$

由Ω的一切子集构成事件域F，它包含$\binom{8}{0}+\binom{8}{1}+\cdots+\binom{8}{8}$ $=2^8$个元素。今设“两次正面一次反面”的事件为A，即

$$\begin{aligned}A&=\{A_1,A_2,A_3\}\\&=\{001,010,100\}\in F.\end{aligned}$$

如果硬币是匀称的，即掷一次出现正面与反面的概率都是$\frac{1}{2}$，则有

$$P(A_1)=\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{8},\quad P(A_2)=\frac{1}{8},\quad P(A_3)=\frac{1}{8},$$

因此，

$$P(A)=\sum_{i=1}^{3}P(A_i)=\frac{3}{8}.$$

如果硬币不匀称，设出现正面的概率为p，出现反面的概率是$1-p$，则有

$$P(A)=\binom{3}{2}p^2(1-p).$$

由定义容易推出概率P具有下列性质：

(1) 不可能事件的概率为零，即

$$P(\overline{\Omega})=P(\varnothing)=0. \tag{1·1·2}$$

(2) 有限可加性，即

若事件$A_i\cap A_j=\varnothing$ $(i\neq j)$，则

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)\cup P(A_2)\cup\cdots\cup P(A_n). \tag{1·1·3}$$

(3) 对任何事件$A\in F$，恒有

$$P(\overline{A})=1-P(A). \tag{1·1·4}$$

(4) 如果事件A、$B\in F$，且$A\supset B$，则

$$P(A-B)=P(A)-P(B),$$

且

$$P(A)\geqslant P(B). \tag{1·1·5}$$

(5) $P(A\cup B)=P(A)+P(B)-P(A\cap B)$,

且

$$P(A\cup B)\leqslant P(A)+P(B). \tag{1·1·6}$$

证明从略。

§1·2 随机变量及其概率分布

1. 随机变量

在随机试验中，试验的结果常常可用一个数量来刻划它。例如要检查一批产品的质量，需要对该批产品进行随机抽样，如果在一次抽样中出现废品的个数是ξ，我们关心的是“$\xi=a$”($a=0,1,2,\cdots$)或“$\xi<a$”等事件的概率。又如在测量中，一次测量的误差大小是随机的，也可用数ξ来表示，我们关心的是数ξ落在区间(a,b)内的概率。

显然，这种数ξ随着试验结果的不同而变化，因而它是样本点ω的函数：$\xi=\xi(\omega)$。由于一次试验出现哪个样本点是随机的，故$\xi(\omega)$的取值也是随机的，因此，称$\xi=\xi(\omega)$为**随机变量**。

如果随机变量ξ的所有可能取值为有限或无限可列个值，则称$\xi=\xi(\omega)$为**离散型随机变量**（如前例中的废品数）。对定义在样本空间Ω上的离散型随机变量$\xi(\omega)$，为使事件$\xi(\omega)=x_i(i=1,2,\cdots)$的概率有意义，$\xi(\omega)=x_i$必须是事件域$F$中的元素，即

$$\{\omega|\xi(\omega)=x_i(i=1,2,\cdots)\}\in F,$$

其中x_i是任意实数。

如果随机变量$\xi=\xi(\omega)$可取某个区间(a,b)或$(-\infty,+\infty)$的一切值（如后例中的误差范围），则称$\xi(\omega)$为**连续型随机变量**。对定义在样本空间Ω上的连续型随机变量$\xi(\omega)$，若要求事件$(a\leqslant\xi(\omega)<b)$或事件$\xi(\omega)\in B$（其中$B$是由区间经和、积运算而得到的一维波雷耳集）的概率有意义，则必须要求事件

$$\{\omega|a\leqslant\xi(\omega)<b\}\in F,$$

或

$$\{\omega|\xi(\omega)\in B\}\in F,$$

下面我们给出随机变量的严格定义：

对给定的概率空间(Ω,F,P)，若$\xi=\xi(\omega)$是样本空间Ω上的一个实值函数，且对任意实数x，事件$(\xi<x)=\{\omega|\xi(\omega)<x\}\in F$，则称$\xi=\xi(\omega)$是$(\Omega,F,P)$上的一个随机变量。

2. 分布函数

定义 对任意实数x，事件$(\xi<x)$的概率$P(\xi<x)$（显然依赖于x）是x的函数，即

$$F(x)=P(\xi<x), \tag{1·2·1}$$

则称$F(x)$为随机变量$\xi=\xi(\omega)$的**分布函数**。

由定义容易推出$F(x)$具有下列性质：

(1) 单调性 $F(x)$是x的不降函数，即

若 $a\leqslant b$，则$F(a)\leqslant F(b)$。

(2) 有界性 $0\leqslant F(x)\leqslant 1$。

$$\lim_{x\to-\infty}F(x)=0,\quad \lim_{x\to\infty}F(x)=1.$$

(3) 左连续性 $F(x-0)=F(x)$。