時間序列分析
—建模與預報

SHI JIAN XU LIE FEN XI JIAN MO YU YU BAO

杜金觀　項靜怡　戴儉華　編著

安徽教育出版社

時間序列分析
——建模與預報

杜金觀　項靜怡　戴儉華　編著

安徽教育出版社

时间序列分析——建模与预报

安徽教育出版社出版

（合肥市金寨路283号）

安徽省新华书店发行　　安徽新华印刷厂印刷

开本850×1168　1/32　印张12.25　字数200,000

1991年5月第1版　　1991年5月第1次印刷

印数4,000

ISBN 7—5336—0632—9/G·1105

定价：5.00元

编 者 的 话

所谓时间序列，又称动态数据，指的是一组按时间顺序排列的数字序列。例如，某类物价指数的波动记录，某地区的年降雨量序列，电机噪声信号序列，机械系统的振动序列等等。这些数据由于偶然因素的影响，往往表现出随机性，相互之间存在某种统计上的联系。

单变量时间序列分析源远流长，但由于计算十分困难复杂，发展缓慢。随着计算技术和计算机的发展，1968年，Box和Jenkins提出一套比较完善的时间序列的建模理论和分析方法，此后，便迅速发展起来。近10多年，它在工业自动化、机械、水文、气象、地质、化工、交通等自然科学和工程技术领域中，以及经济管理、军事科学和某些社会科学领域中都得到了广泛的应用。

时间序列分析有时域分析和频域分析两大部分。本书主要介绍时域和频域分析的基本理论和方法，并侧重于应用。同时，本书中某些内容是近年来发展起来的新成果和新方法。本书是时间序列分析内容较完整的教科书或参考书。全书分为九章：第一章是预备知识，简要地向读者介绍随机序列、统计估计和线性差分方程知识；第二、三、四章是核心部分，叙述单变量时间序列的几类模型及其模型识别、参数估计、模型检验和建模方法；第五章为时间序列预报；第六章是基于多值状态的时间序列分析，主要介绍基于两种状态分析处理动态数据的简便有趣的手段和方法；第七章研究多维时间序列的AR模型识别、参数估计和预报问题；第八章提出了近几年来许多人开始研究的非线性时间序列模型，并着重介绍了在应用上颇有成效的门限自回归模型，采用逐段线性化手段来处理困难的非线性系统；最后一章介绍了时间序列频域分析的基本理论和方法。各章均配有一定数量的习题，

并在附录中给出较详细的解答或解题提示，以方便读者。

本书前五章是基本部分，教学时数约40学时，后四章打星号“*”，说明这个部分内容是为更高要求的专业和读者编写的，讲完全书约80学时。

本书出版之前，曾在全国工科院校应用概率统计委员会举办的讲习班上，由编者多次讲授过，并在合肥工业大学等院校作为工科研究生和本科生教材试用过，从而有机会得到较多的学者的热心帮助，在此向他们表示衷心的感谢。尽管如此，限于作者水平和经验，书中肯定有不妥之处，谨希读者不吝指正，使它能逐步成为一本较畅销的教科书或参考书。

编者

1990年1月

符 号 说 明

$\triangleq$	定义为
A^{τ}	矩阵A的转置
$\mathrm{tr}A$	矩阵A的迹
$\{X_K\}$	随机序列
$\{\mathbf{X}\}$	随机向量序列
$\hat{\mathbf{X}}(\mathbf{Z})$	基于量测$\mathbf{Z}$的对$\mathbf{X}$的某种估计
$\tilde{\mathbf{X}}$	估计的误差
$\mathbf{X}_{ML}$	随机向量$\mathbf{X}$的极大似然估计
$\mathbf{X}_{LS}$	随机向量$\mathbf{X}$的最小二乘估计
$\mathbf{X}_{MV}$	随机向量$\mathbf{X}$的最小方差估计
$\hat{\boldsymbol{\beta}}$	参数向量$\boldsymbol{\beta}$的估计
$\tilde{\boldsymbol{\beta}}$	参数向量估计的误差
$\hat{x}_k(l)$	基于量测…，x_1，x_2，…，x_k对$x_{k+l}(l>0)$的平稳预报
$\hat{x}_{k\mid j}$	基于量测x_1，x_2，…，$x_j(j<k)$对x_k的平稳预报
B^k	k步延迟算子（后移算子）
$\ll$	远小于
$\gg$	远大于
SUP	上确界
$\cong$	近似等于
$\approx$	渐近等于
$\xrightarrow{p}$	依概率收敛

$\xrightarrow{a\cdot s}$　概率为1地收敛

$\mathrm{l\cdot i\cdot m}$　均方极限

目　录

第一章　预备知识

§1　随机序列

§1.1　随机序列的定义

在概率论和数理统计中仅仅涉及到一个随机变量X或一个随机向量$\mathbf{X}$。而时间序列分析所考察的对象，常常不是单个随机变量或单个随机向量，而是一族随机变量或一族随机向量。例如，X_t表示t时刻的气温，显然是一个随机变量，当t在某段时间$[a, b]$内考察的话，便得到一族随机变量为$\{X_t;\ t\in[a,\ b]\}$。又如X_t表示某商场在t时刻的营业额，将t从某年某月某时（以小时为单位）开始记录下来，就可得到一列随机变量。因此，一般地需要引进随机过程和随机序列的概念。

设T是某个集合，俗称足标集，对任意固定$t\in T$，X_t是随机变量，当t在T中跑遍时，得到随机变量的全体$\{X_t:t\in T\}$，记为X_T或$\{X_t\}$，称X_T为T上的随机函数。通常T取为

$$1^\circ\ T=(-\infty,\ +\infty),\ T=[0,\ +\infty)$$

$$2^\circ\ T=\{\cdots,\ -2,\ -1,\ 0,\ 1,\ 2,\ \cdots\},\ T=\{1,\ 2,\cdots\}$$

随机函数X_T中T取1°的情形，称X_T为随机过程；T取2°的情形，称X_T为随机序列，后者有时简记为$\{X_k\}$，许多实际问题，随机序列是从随机过程按某一采样间隔Δ得到，故2°的T实际上是$T=\{\cdots,\ -2\Delta,\ -1\Delta,\ 0,\ \Delta,\ 2\Delta,\ \cdots\}$等等。

从概率统计便知，若X为某一试验结果的随机变量，那么每

做一次试验，就能获得X的一个取值x，称x为X的一个样本值。对于随机过程X_T而言，它的一个样本函数x_T是T上的函数，在1°的情形就是一条曲线或称轨迹，在很多实际情形，可由自动记录仪记录下来，有时常称这函数为随机过程X_T的一个现实。例如，信息传输中接收机的噪声电压就是随时间变化的随机变量，即随机过程X_T。当我们对接收机的输出电压(或电流)作单次观察时，可能得到图1-1的起伏波形$x_1(t)$。实际上，在试验结果中出现的噪声电压的具体波形也可能是另外的样子，如$x_2(t)$，$x_3(t)$，等等。所有这些以一定概率的可能的$x_1(t)$，$x_2(t)$，$x_3(t)\cdots$的集合就构成了随机过程X_T，而$x_1(t)$，$x_2(t)$，$x_3(i)\cdots$等等都是该随机过程的一个个样本函数或现实，它们都是确定的时间函数。

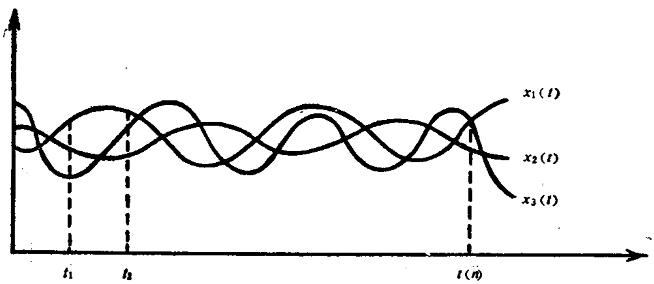

图 1-1

由图1-1可以看出，当参变量t固定在某一时刻t_1时，各个样本函数的取值$x_1(t_1)$，$x_2(t_1)$，$x_3(t_1)$，…等等的大小是各个不相同的值，这时随机过程X_T就是一般意义下的随机变量X_{t_1}。同样，当t取不同的固定值t_2，t_3，…t_n时，随机过程就是一族随机变量。

对于随机序列而言，样本函数$\{x_k\}$是一普通实数列，称$\{x_k\}$为随机序列$\{X_k\}$的一个现实。在实际问题中，因为随时间t的流逝不能重复，所以我们往往仅能获得随机序列(或过程)的一个现

实，而长度为N的动态数x_1，x_2，…，x_N常常是随机过程X_T按一定的采样间隔Δ而获得的样本值，今后称x_1，x_2，…，x_N为随机序列$\{x_K\}$的长度为N的样本值，有时作为长度为N的样本时，以大写X_1，X_2。…X_N表示，但经常大写X_k与小写x_k不加区别，读者根据上下文会自明。

另外，随机序列$\{X_k\}$的整数变量k，通常表示采样间隔Δ的k倍，例如$k\Delta$表示第k小时，第k天，生物学上表示第k代，地质学上表示第k层等，但大量的问题与时间有关，所以，常称$\{X_k\}$为时间序列。

§1.2 随机序列的概率分布

一个随机变量的统计特性完全由它的分布函数$F(x)$或分布密度$p(x)$所确定。同样，一个随机向量$\mathbf{X}=(X_1, X_2, \cdots, X_n)^\tau$的统计特性完全由它的联合分布函数$F(x_1, x_2, \cdots, x_n)$或联合分布密度$p(x_1, x_2, \cdots, x_n)$所确定，其中$\tau$表示转置。对于随机序列而言，由于它是可列个随机变量构成的。因此，对于任意$t\in T$，X_t由分布函数$F_t(x)$来描述，对于任意t_1，$t_2\in T$，$(X_{t1}, X_{t2})^\tau$由联合分布函数$F_{t_1, t_2}(x_1, x_2)$来描述，…，对于任意正整数n和任意t_1，t_2，…$t_n\in T$，$(X_{t1}, X_{t2}, \cdots, X_{tn})$由联合分布函数来描述，其中$F_{t_1, \cdots, t_n}(x_1, \cdots, x_n)=P(X_{t1}<x_1, \cdots X_{tn}<x_n)$。(1.1.1)

这样得到的任意正整数k相应k维联合分布族：

$\{F_{t1, \cdots, tk}(x_1, \cdots x_k)$：任意正整数$k$；$t_1$，…$tk\in T$，$k=1.2, \cdots\}$，我们称上述分布族为随机序列$\{X_k\}$的有穷维分布族。

如果随机序列$\{Xk\}$任意有穷维分布满足

$$F_{t1, t2, \cdots, tn}(x_1, x_2, \cdots, x_n)=F_{t1}(x_1)F_{t2}(x_2)\cdots F_{tn}(x_n) \tag{1.1.2}$$

即序列中任意个随机变量X_{t1}，X_{t2}，…X_{tn}都相互独立，则称

$\{X_k\}$ 为独立随机序列，实际应用中，常见的电路中热噪声往往近似于独立随机序列。

随机序列的有穷维分布函数族能完善地刻划随机序列的统计特性，但是在实际应用中想要确定一个随机序列的分布函数族却十分困难，甚至不可能办到。此外，即使知道了这些分布函数族，也不便于分析和实际应用。因此，有必要象引入随机变量的数字特征那样，引入随机序列(或过程)的基本参数特征。当然，这些参数特征要求一方面能描述随机序列的特征，另一方面又要便于测量和计算。

下面将引入随机序列或随机过程的基本参数特征：均值函数、自协方差函数、自相关函数等。

§1.3 随机序列的参数特征

如果对于每个k而言，二阶原点矩$EX_k^2<+\infty$，称$\{X_k\}$为二阶矩有穷的随机序列。对二阶矩有穷的随机序列可定义它的均值函数、自协方差函数和自相关函数。

1.均值函数

$\{X_t\}$为随机序列，称$EX_t=\mu_t$为X_t的均值，称$\{\mu_t\}$为$\{X_t\}$的均值函数。若X_t的分布函数为$F_t(x)$，分布密度为$p_t(x)$，则

$$\mu_t=\int_{-\infty}^{+\infty}xdF_t(x)=\int_{-\infty}^{+\infty}xp(t)dx \qquad (1.1.3)$$

2.自协方差函数

为了分析随机序列$\{X_i\}$中在不同时刻随机变量之间的统计关系，需要对任意不同的整数t，s，考虑X_t与X_s的相互关系。令

$$\begin{aligned}\gamma_{ts}&=E(X_t-\mu_t)(X_s-\mu_s)\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x_1-\mu_t)(x_2-\mu_s)aF_{t,s}(x_1,\ x_2)\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x_1-\mu_t)(x_2-\mu_s)p_{t,s}(x_1,x_2)dx_1dx_2 \qquad (1.1.4)\end{aligned}$$

其中 $F_{t,s}(x_1, x_2)$和$p_{t,s}(x_1, x_2)$ 分别是X_t与X_s 的联合分布函数和分布密度，称 $\{\gamma_{ts}\}$为随机序列 $\{X_t\}$ 的自协方差函数，今后记为$\gamma_{ts}=\text{cov}(X_t, X_s)$，显然$\gamma_{ts}$ 是二元对称函数。特别当 $t=s$时，$\gamma_{tt}=E(X_t-\mu_t)^2$，称 $\{\gamma_{tt}\}$ 为随机序列 $\{X_t\}$ 的方差函数，记为$\gamma_{tt}=\text{var}X_t$

3.自相关函数

设$\{\gamma_{ts}\}$为随机序列$\{X_t\}$的自协方差函数，令

$$\rho_{ts}=\frac{\gamma_{ts}}{\sqrt{\gamma_{tt}}\sqrt{\gamma_{ss}}}=\frac{\text{cov}(X_t, X_s)}{\sqrt{\text{var}X_t\cdot\text{var}X_s}} \tag{1.1.5}$$

称$\{\rho_{ts}\}$为随机序列$\{X_t\}$的自相关函数，这里ρ_{ts}是无量纲并依赖t，s的，它同γ_{ts}一样刻划了$\{X_t\}$中不同时刻t与s的随机变量X_t与X_s的统计相关程度。

若随机序列$\{X_t\}$满足：对任意整数t, s，当$t\neq s$时，$\rho_{ts}=0$，则称$\{X_t\}$为不相关随机序列。

由柯西—希瓦兹（Cauchy—Schwarz）不等式，可以证明对于任意 t，s，ρ_{ts}具有下列性质：

1） $|\rho_{ts}|\leqslant 1$

2） $|\rho_{ts}|=1$ 的充分必要条件是X_t与X_s依概率1线性相关，即$p\{X_s=aX_t+b\}=1$，其中$a\neq 0$，b为常数。

显然，$\{\mu_t\}$，$\{\gamma_{ts}\}$，$\{\rho_{ts}\}$都由随机序列$\{X_t\}$的一维和二维分布族唯一确定，但是由它们都不能确定$\{X_t\}$的有穷维分布族。从理论角度看，仅仅研究均值函数和自协方差函数或自相关函数当然是不能代替整个随机序列的研究，但由于它们刻划了随机序列的主要特性，而且易于测算，故能起到极为重要的作用。

§2 平稳随机序列

平稳随机过程或平稳随机序列，是一类应用极为广泛的随机

过程和随机序列。它们的基本特点是其统计特性不随时间的平移而变化，或者说不随时间原点的选取而变化，在物理上通常当一个稳定系统达到稳态时所产生的过程，因而这种状态有时也称为“统计平衡”状态。

§2.1 严平稳随机序列定义

定义1 设$\{X_t\}$，$t=\cdots, -1, 0, 1, \cdots$为随机序列，任取$t_1, t_2\cdots, t_n\in\{\cdots, -1, 0, 1\cdots\}$，$X_{t_1}, X_{t_2}, \cdots X_{t_n}$的联合分布函数具有对时间推移不变的性质：

$$F_{t_1, t_2\cdots, t_n}(x_1, x_2\cdots x_n)=F_{t_1+\tau, t_2+\tau\cdots, t_n+\tau}(x_1, x_2\cdots, x_n)$$

其中τ为任意整数，则称$\{X_t\}$为严平稳随机序列或强平稳序列。

严平稳随机序列具有下列特性：

1. 一维分布函数与时间无关，即

$$F_t(x)=F_{t+\tau}(x)=F(x) \tag{1.2.1}$$

2. 二维分布函数与时间间隔有关，而同时间起点无关，即取$\tau=-t_2$

$$F_{t_1, t_2}(x_1, x_2)=F_{t_1+\tau, t_2+\tau}(x_1, x_2)$$
$$=F_{t_1-t_2, 0}(x_1, x_2) \tag{1.2.2}$$

3. (i)若$\{x_t\}$为二阶矩有穷的严平稳序列，则均值与方差函数是同t无关的常数，即

$$\mu_t=EX_t=\mu \tag{1.2.3}$$

$$\gamma_{tt}=E(X_t-\mu_t)^2=\sigma^2 \tag{1.2.4}$$

(ii)自协方差与自相关函数仅同时间间隔$\tau=t_2-t_1$有关，即

$$\gamma_{t_1t_2}=E(X_{t_1}-\mu)(X_{t_2}-\mu)=\gamma_\tau \tag{1.2.5}$$

$$\rho_{t_1t_2}=\frac{\gamma_{t_1t_2}}{\sqrt{\gamma_{t_1t_1}\cdot\gamma_{t_2t_2}}}=\rho_\tau \tag{1.2.6}$$

§2.2 宽平稳随机序列

严平稳随机序列的要求十分苛刻。通常时间序列分析所研究的是所谓宽平稳随机序列。

定义2 如果随机序列$\{X_t\}$是二阶矩有穷，即对任何整数t，有$EX_t^2<\infty$，且满足下列条件：

1. 对任意整数t，$EX_t=\mu$，μ为常数；

2. 对任意整数t，s，自协方差函数γ_{ts}仅与$t-s$有关，而同个别时刻t，s无关，即

$$\gamma_{ts}=\gamma_{t-s}$$

则称$\{X_t\}$为宽平稳随机序列或弱平稳随机序列，简称为平稳随机序列。

今后若无特别说明，所谓平稳随机序列指的是宽平稳随机序列。

注意，严平稳随机序列不一定是宽平稳随机序列。例如，$\{X_k\}$，$k=0$，± 1，…是独立随机序列，且X_k服从柯西分布，即分布密度为

$$p(x)=\frac{1}{\pi(1+x^2)}$$

显然$\{X_k\}$是严平稳的，但服从柯西分布的随机变量X_k的均值不存在，从而X_k的二阶矩不是有穷，故$\{X_k\}$不是宽平稳随机序列。反之，宽平稳随机序列也未必是严平稳随机序列。例如，设$X_k=\sin 2k\pi\theta$，$k=1,2,\cdots$其中θ是〔0，1〕区间上均匀分布随机变量，显然

$$EX_k=\int_0^1 \sin 2k\pi x dx=0 \qquad k=1,2,\cdots$$

$$EX_kX_l=\int_0^1 \sin 2k\pi x \sin 2l\pi x dx=\frac{1}{2}\delta_{kL}$$