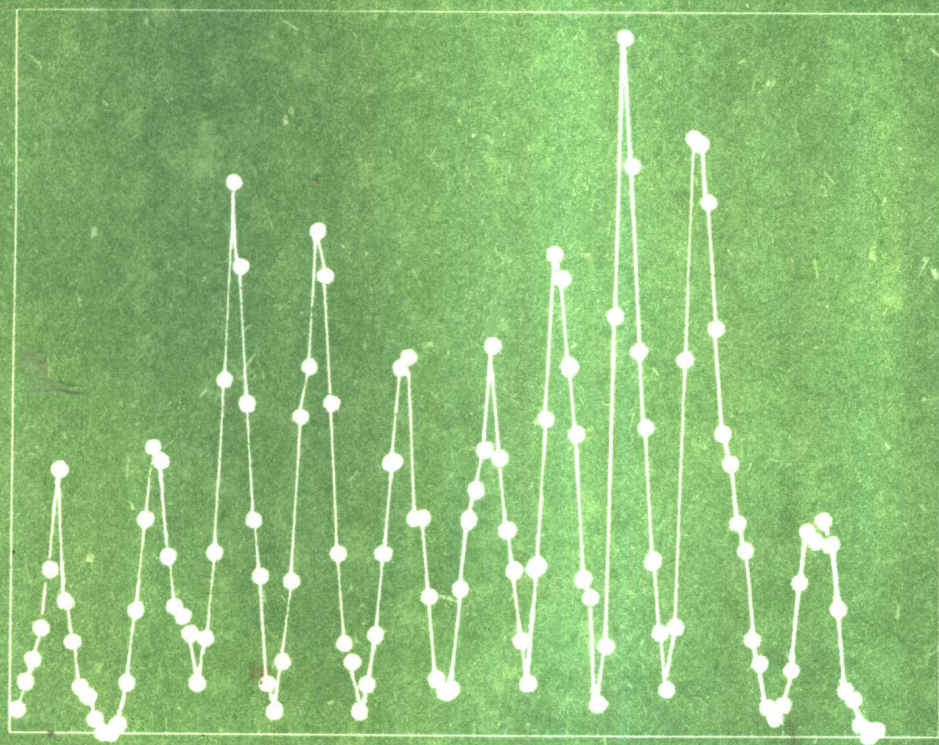


高等学校教材

# 动态数据处理的理论与方法

## ——时间序列分析

田 铮 主编



西北工业大学出版社

高等学校教材

# 动态数据处理的理论与方法

——时间序列分析

田 铮 主编

田 铮 戎海武 编

西北工业大学出版社

1995年6月 西安

(陕)新登字 009 号

【内容简介】 本书从时域分析和频域分析两个角度,系统地介绍了时间序列分析的基本理论和方法。书中主要内容有平稳时间序列的基本概念;Hilbert 空间的有关理论;ARMA 模型;时间序列的预报;时间序列的谱表示和谐分析的初步理论;非平稳时间序列的 ARIMA 模型及多维时间序列的初步理论等 7 个部分组成。

本书可作为工科院校有关专业和应用数学专业的高年级学生和研究生选修课教材,也可供从事动态数据分析与处理的科技工作者参考。

高等学校教材  
动态数据处理的理论与方法  
——时间序列分析

主 编 田 铮  
责任编辑 刘彦信  
责任校对 樊 力

©1995 西北工业大学出版社出版  
(710072 西安市友谊西路 127 号 电话 4253407)  
陕西省新华书店发行  
西北工业大学出版社印刷厂印装  
ISBN 7-5612-0745-X/O·100(课)

开本 787×1092 毫米 1/16 11.625 印张 字数 281 千字  
1995 年 6 月第 1 版 1995 年 6 月第 1 次印刷  
印数:1—1 000 册 定价:10.00 元

购买本社出版的图书,如有缺页、错页的,本社发行部负责调换。

## 前 言

时间序列分析是概率统计学的一个重要的新分支。它是用概率统计方法研究随时间变化、具有随机性而又前后相互关联的动态数据(称其为随机时间序列,简称为时间序列)。时间序列分析的研究内容包括时域分析和频域分析,其中包括建立模型、参数估计、最佳预测和控制以及谱估计等理论和方法。近年来,时间序列分析已在控制、人工神经网络、生物、医学、水文、气象、经济学等自然科学和社会科学领域中得到广泛的应用。

自1988年以来,编者为西北工业大学应用数学系本科生和硕士研究生分别开设了“时间序列分析”课程,积累了比较丰富的经验和资料。本书就是在为该课程编写并经过多次教学使用(其中包括四届本科生和四届研究生)的讲义的基础上,吸取了目前国内外同类书籍长处,经修改、充实编写而成的。

本书注重基本理论、基本概念和基本方法的论述,力求做到论述严谨清晰,文字易懂,数学工具的应用简明、准确,以便读者系统掌握动态数据处理的基本理论和方法,进而了解和掌握时间序列分析发展的一些新内容。阅读和学习本书须掌握概率论、线性代数和泛函分析的有关基础知识。

本书取材较为全面,内容包括一维平稳时间序列的时域分析和谱分析的基本理论和方法,时间序列预报的时域方法和频域方法,动态数据的模型识别以及非平稳时间序列模型和多维时间序列分析初步。为方便读者阅读和学习,书后附录一介绍了常系数线性差分方程,附录二介绍了随机变量序列的收敛性和中心极限定理。

本书的第一章—第八章和附录二由田铮编写,第九章和附录一由戎海武编写,全书由田铮主编。

西安电子科技大学博士导师肖国镇教授审阅了全部书稿,并提出许多宝贵意见;本书的出版得到西北工业大学出版社、西北工业大学应用数学系的热情支持和帮助,谨此一并致谢。

书中的错误之处恳请读者批评指正。

田 铮

1994年8月

# 目 录

|                                                            |    |
|------------------------------------------------------------|----|
| 第一章 平稳时间序列.....                                            | 1  |
| § 1.1 时间序列的实例 .....                                        | 1  |
| § 1.2 随机过程的基本概念 .....                                      | 7  |
| § 1.3 平稳过程的自协方差函数.....                                     | 12 |
| § 1.4 趋势项和季节项的估计和分离.....                                   | 14 |
| § 1.5 Kolmogorov 定理的应用 .....                               | 18 |
| 习题一 .....                                                  | 19 |
| 第二章 Hilbert 空间和 $L^2(\Omega, \mathcal{F}, P)$ 空间中的预报 ..... | 21 |
| § 2.1 内积空间及其性质.....                                        | 21 |
| § 2.2 Hilbert 空间和预报方程 .....                                | 23 |
| § 2.3 $L^2(\Omega, \mathcal{F}, P)$ 空间中的预报 .....           | 32 |
| § 2.4 Fourier 级数 .....                                     | 33 |
| § 2.5 Hilbert 空间的同构 .....                                  | 38 |
| § 2.6 $L^2(\Omega, \mathcal{F}, P)$ 空间的完备性 .....           | 39 |
| 习题二 .....                                                  | 39 |
| 第三章 ARMA 序列.....                                           | 42 |
| § 3.1 因果可逆的 ARMA 序列 .....                                  | 42 |
| § 3.2 ARMA 序列的自相关(系数)函数 .....                              | 50 |
| § 3.3 ARMA 序列的偏相关(系数)函数 .....                              | 57 |
| 习题三 .....                                                  | 62 |
| 第四章 时间序列的谱表示 .....                                         | 66 |
| § 4.1 复值平稳时间序列.....                                        | 66 |
| § 4.2 平稳序列自协方差函数的谱表示.....                                  | 66 |
| § 4.3 ARMA 序列的有理谱密度 .....                                  | 70 |
| § 4.4 平稳序列的谱表示.....                                        | 76 |
| 习题四 .....                                                  | 84 |
| 第五章 时间序列的预报 .....                                          | 87 |
| § 5.1 时间序列时域中的预报方法.....                                    | 87 |
| § 5.2 最佳线性预报的递推算法.....                                     | 89 |
| § 5.3 ARMA( $p, q$ ) 序列的递推预报 .....                         | 94 |

|                                                   |     |
|---------------------------------------------------|-----|
| § 5.4 时间序列预报的频域方法 .....                           | 100 |
| 习题五 .....                                         | 101 |
| <b>第六章 平稳时间序列的模型拟合</b> .....                      | 103 |
| § 6.1 平稳序列样本均值和样本自协方差函数的统计性质 .....                | 103 |
| § 6.2 $AR(p)$ 模型的参数估计 .....                       | 107 |
| § 6.3 $MA(q)$ 模型的参数估计 .....                       | 110 |
| § 6.4 $ARMA(p, q)$ 模型的参数估计 .....                  | 114 |
| 习题六 .....                                         | 118 |
| <b>第七章 非平稳时间序列的 ARIMA 模型和时间序列的模型识别</b> .....      | 120 |
| § 7.1 非平稳时间序列的 ARIMA 模型 .....                     | 120 |
| § 7.2 季节 ARIMA 模型 .....                           | 124 |
| § 7.3 模型阶数的估计 .....                               | 130 |
| § 7.4 时间序列的统计检验和模型拟合优度检验 .....                    | 131 |
| § 7.5 时间序列的模型识别 .....                             | 135 |
| 习题七 .....                                         | 142 |
| <b>第八章 时间序列的谱分析初步</b> .....                       | 144 |
| § 8.1 周期图 .....                                   | 144 |
| § 8.2 隐含周期项的检验 .....                              | 149 |
| § 8.3 ARMA 谱估计 .....                              | 154 |
| 习题八 .....                                         | 157 |
| <b>第九章 多维时间序列分析初步</b> .....                       | 158 |
| § 9.1 多维平稳时间序列的二阶矩性质 .....                        | 158 |
| § 9.2 多维 ARMA 序列 .....                            | 160 |
| § 9.3 二阶矩存在的随机向量的最佳线性预报和<br>多维 ARMA 序列的递推预报 ..... | 162 |
| 习题九 .....                                         | 166 |
| <b>附录</b> .....                                   | 167 |
| 附录一 常系数线性差分方程 .....                               | 167 |
| 附录二 随机变量序列的收敛性和中心极限定理 .....                       | 169 |
| <b>参考文献</b> .....                                 | 180 |

# 第一章 平稳时间序列

时间序列是依时间变化而又相互关联的数据序列,对时间序列进行统计分析,简称为时间序列分析。本章介绍时间序列和随机过程的基本概念以及由动态数据中估计和分离趋势项和周期项的方法。

## § 1.1 时间序列的实例

通常按时间顺序排列的观测数据序列 $\{x_t, t \in T_0\}$ 称为一个时间序列,其中 $T_0$ 为指标集。如果 $T_0$ 是离散集,则称时间序列 $\{x_t, t \in T_0\}$ 为离散时间序列;如果 $T_0$ 为连续集,则称其为连续时间序列,记为 $\{x(t), t \in T_0\}$ 。本书主要研究离散时间序列。如果在指标集 $T_0$ 的某个时刻同时观测到多个数据序列,则称 $\{x_t, t \in T_0\}$ 为多维时间序列。本书以介绍一维时间序列的理论和方法为主。值得注意的是指标集 $T_0$ 并不总是时间集。

### 一、时间序列的实例

例 1.1.1—例 1.1.10 给出了 10 个时间序列实例,为简便,将其相应地称为序列 1—序列 10。

例 1.1.1 某地区 1790—1980 年人口总数序列(表 1.1,图 1.1)。

表 1.1 某地区人口总数  $x_t$  (1790—1980 年)

| $t$ (年) | $x_t$ (人)  | $t$ (年) | $x_t$ (人)   |
|---------|------------|---------|-------------|
| 1790    | 3 929 214  | 1890    | 62 979 766  |
| 1800    | 5 308 483  | 1900    | 76 212 168  |
| 1810    | 7 239 881  | 1910    | 92 228 496  |
| 1820    | 9 638 453  | 1920    | 106 021 537 |
| 1830    | 12 860 702 | 1930    | 123 202 624 |
| 1840    | 17 063 353 | 1940    | 132 164 569 |
| 1850    | 23 191 876 | 1950    | 151 325 798 |
| 1860    | 31 443 321 | 1960    | 179 323 175 |
| 1870    | 38 558 371 | 1970    | 203 302 031 |
| 1880    | 50 189 209 | 1980    | 226 545 805 |

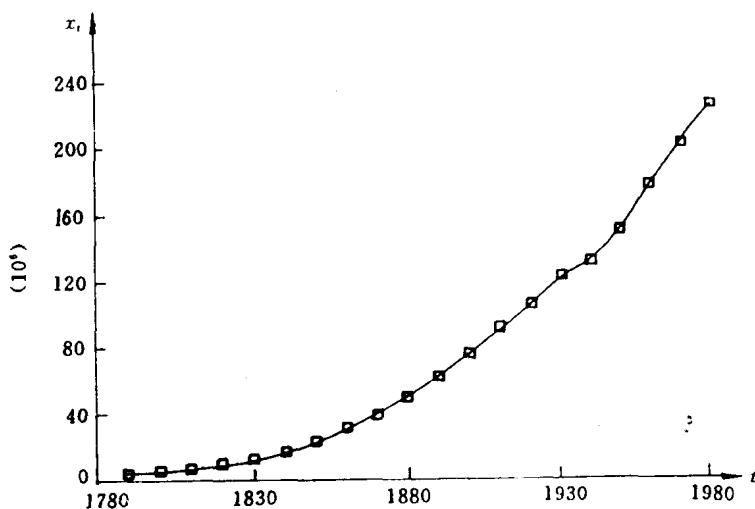


图 1.1 某地区人口总数  $x_t$  (1780—1980 年)

例 1.1.2 (通过电阻的电流) 电阻  $r$  两端加一正弦波电压  $v(t) = a\cos(Vt + \theta)$ , 连续记录下电流(图 1.2), 则得一连续时间序列

$$x(t) = r^{-1}a\cos(Vt + \theta)$$

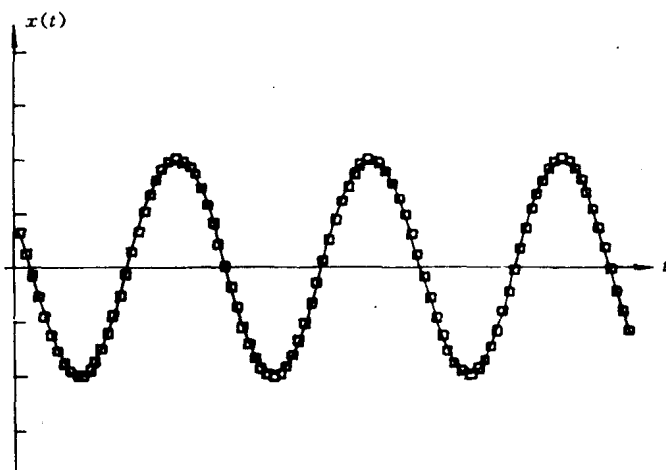


图 1.2 序列  $x(t) = \cos(0.2t + \frac{\pi}{3})$  的观测数据

例 1.1.3 白噪声序列(图 1.3)。

例 1.1.4 某地区降雨量数据(图 1.4) 序列  $x_t$  (1817—1922 年)。

例 1.1.5 随机游动序列(图 1.5)。

例 1.1.6 某湖水位数据(图 1.6) 序列  $x_t$  (1921—1971 年)。

例 1.1.7 Wölfer 太阳黑子数(表 1.2, 图 1.7)  $x_t$  (1770—1869 年)。

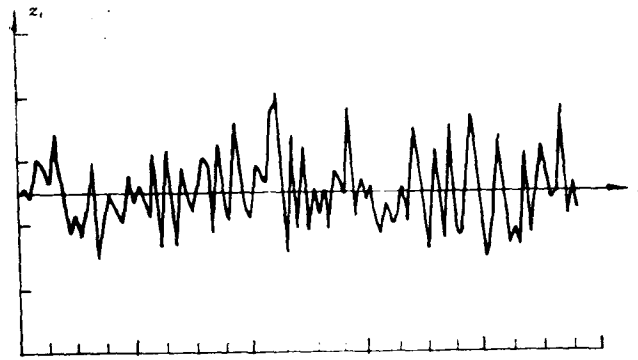


图 1.3 白噪声序列

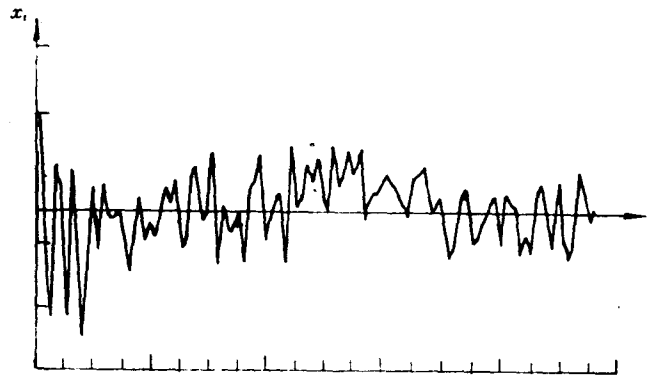


图 1.4 某地区降雨量数据  $x_t$  (1817 — 1922 年)

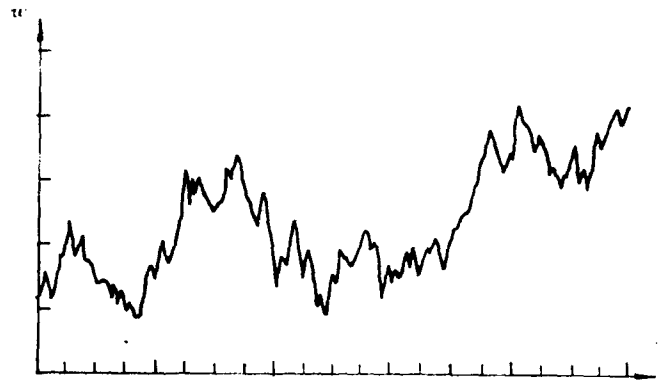


图 1.5 随机游动序列

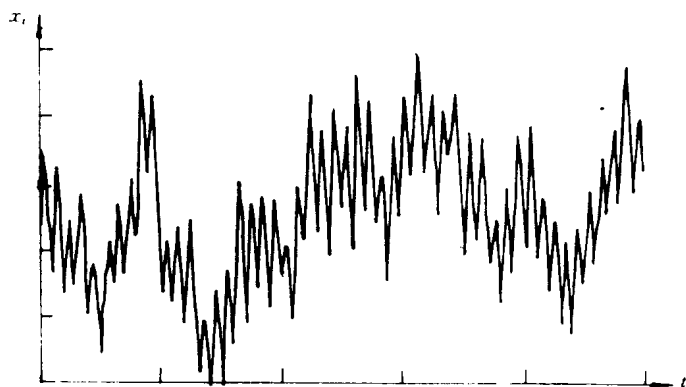


图 1.6 某湖水位数据序列  $x_t$  (1921 — 1971 年)

表 1.2 Wölfer 太阳黑子数  $x_t$  (1770 — 1869 年)

| $t$ (年) | $x_t$ | $t$ (年) | $x_t$ | $t$ (年) | $x_t$ | $t$ (年) | $x_t$ | $t$ (年) | $x_t$ |
|---------|-------|---------|-------|---------|-------|---------|-------|---------|-------|
| 1770    | 101   | 1790    | 90    | 1810    | 0     | 1830    | 71    | 1850    | 66    |
| 1771    | 82    | 1791    | 67    | 1811    | 1     | 1831    | 48    | 1851    | 64    |
| 1772    | 66    | 1792    | 60    | 1812    | 5     | 1832    | 28    | 1852    | 54    |
| 1773    | 35    | 1793    | 47    | 1813    | 12    | 1833    | 8     | 1853    | 39    |
| 1774    | 31    | 1794    | 41    | 1814    | 14    | 1834    | 13    | 1854    | 21    |
| 1775    | 7     | 1795    | 21    | 1815    | 35    | 1835    | 57    | 1855    | 7     |
| 1776    | 20    | 1796    | 16    | 1816    | 46    | 1836    | 122   | 1856    | 4     |
| 1777    | 92    | 1797    | 6     | 1817    | 41    | 1837    | 138   | 1857    | 23    |
| 1778    | 154   | 1798    | 4     | 1818    | 30    | 1838    | 103   | 1858    | 55    |
| 1779    | 125   | 1799    | 7     | 1819    | 24    | 1839    | 86    | 1859    | 94    |
| 1780    | 85    | 1800    | 14    | 1820    | 16    | 1840    | 63    | 1860    | 96    |
| 1781    | 68    | 1801    | 34    | 1821    | 7     | 1841    | 37    | 1861    | 77    |
| 1782    | 38    | 1802    | 45    | 1822    | 4     | 1842    | 24    | 1862    | 59    |
| 1783    | 23    | 1803    | 43    | 1823    | 2     | 1843    | 11    | 1863    | 44    |
| 1784    | 10    | 1804    | 48    | 1824    | 8     | 1844    | 15    | 1864    | 47    |
| 1785    | 24    | 1805    | 42    | 1825    | 17    | 1845    | 40    | 1865    | 30    |
| 1786    | 83    | 1806    | 28    | 1826    | 36    | 1846    | 62    | 1866    | 16    |
| 1787    | 132   | 1807    | 10    | 1827    | 50    | 1847    | 98    | 1867    | 7     |
| 1788    | 131   | 1808    | 8     | 1828    | 62    | 1848    | 124   | 1868    | 37    |
| 1789    | 118   | 1809    | 2     | 1829    | 67    | 1849    | 96    | 1869    | 74    |

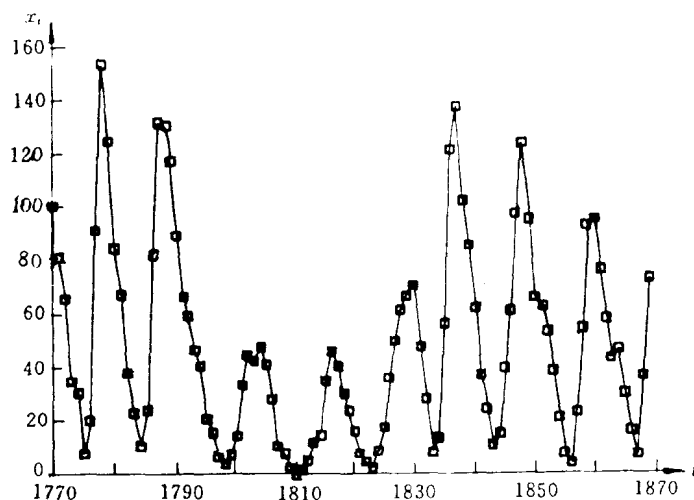


图 1.7 Wöller 太阳黑子数  $x_t$  (1770 — 1869 年)

例 1.1.8 某地区月交通事故死亡数据 (表 1.3, 图 1.8)  $x_t$  (1973 — 1978 年)。

表 1.3 某地区月交通事故死亡数据  $x_t$  (1973 — 1978 年)

| 月 \ 年 | 1973   | 1974   | 1975   | 1976   | 1977   | 1978   |
|-------|--------|--------|--------|--------|--------|--------|
| 1     | 9 007  | 7 750  | 8 162  | 7 717  | 7 792  | 7 836  |
| 2     | 8 106  | 6 981  | 7 306  | 7 461  | 6 957  | 6 892  |
| 3     | 8 928  | 8 038  | 8 124  | 7 776  | 7 726  | 7 791  |
| 4     | 9 137  | 8 422  | 7 870  | 7 925  | 8 106  | 8 129  |
| 5     | 10 017 | 8 714  | 9 387  | 8 634  | 8 890  | 9 115  |
| 6     | 10 826 | 9 512  | 9 556  | 8 945  | 9 299  | 9 434  |
| 7     | 11 317 | 10 120 | 10 093 | 10 078 | 10 625 | 10 484 |
| 8     | 10 744 | 9 823  | 9 620  | 9 179  | 9 302  | 9 827  |
| 9     | 9 713  | 8 743  | 8 285  | 8 037  | 8 314  | 9 110  |
| 10    | 9 938  | 9 129  | 8 433  | 8 488  | 8 850  | 9 070  |
| 11    | 9 161  | 8 710  | 8 160  | 7 874  | 8 265  | 8 633  |
| 12    | 8 927  | 8 680  | 8 034  | 8 647  | 8 796  | 9 240  |

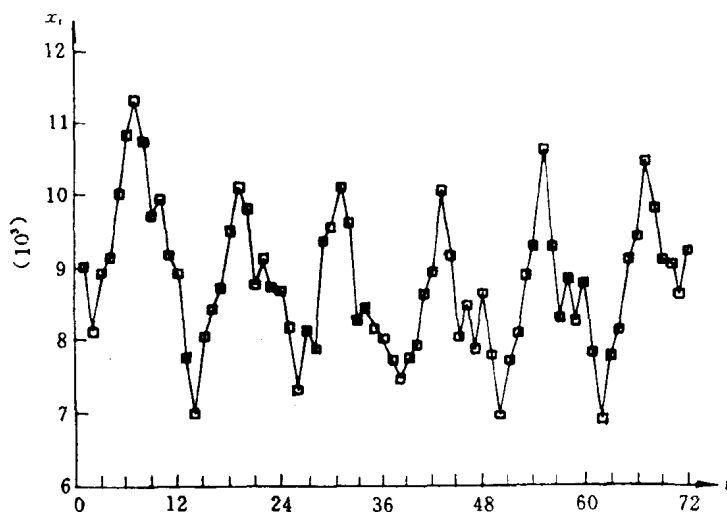


图 1.8 某地区月交通事故死亡数据  $x_t$  (1973 — 1978 年)

例 1.1.9 某航空公司国际航线旅客每月统计(图 1.9, 表 7.1) 数据  $v_t$  (1949 年 1 月 — 1960 年 12 月)。

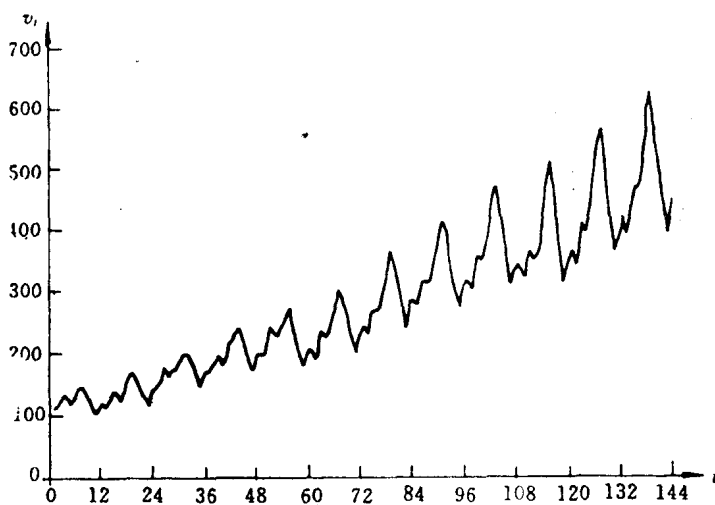


图 1.9 某国际航空公司旅客每月统计数据  $\{v_t, t = 1, \dots, 144\}$   
(1949 年 1 月 — 1960 年 12 月)

例 1.1.10 某一化学反应中输入和输出气体(图 1.10) 数据  $\{x_t, t \in T\}$ 。

由图 1.1 — 图 1.10 可见, 10 个时间序列完全不同。图 1.1 中的曲线是近似于指数递增的曲线; 图 1.7 是一含有周期为 11 的曲线; 图 1.8 是一含有周期为 12 的曲线; 图 1.10 所示是一个二维时间序列。

在数理统计课程中, 介绍过假设检验等基本理论和方法, 但所研究的数据序列彼此间是相互独立的, 是来自某一总体或某些总体的不同独立随机抽样值。而时间序列分析所研究的动态

数据是对同一物理现象在不同时刻(或位置)的重复观测,必须考虑数据间的相关性。从某种意义上可以说,正是这种相关性表征了产生动态数据的物理现象的“记忆”性,同时也使得对时间序列的分析和结论的解释比在独立性条件下要困难得多。

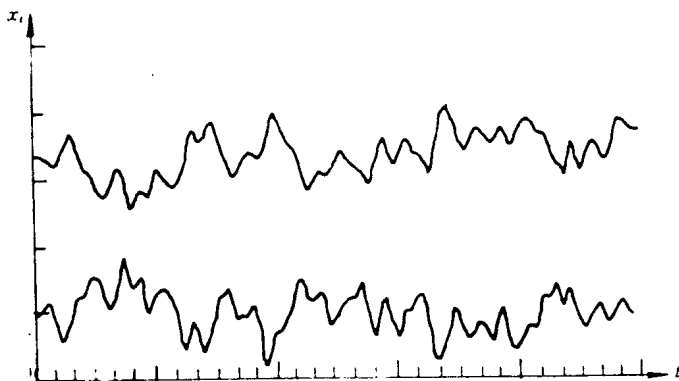


图 1.10 某一化学反应中输入和输出气体数据  $\{x_t, t \in T\}$

## 二、时间序列分析的目的

时间序列分析的目的可以概括为四个方面:描述、推断、预报和控制。描述就是通过对数据的分析建立适当的数学模型来描述产生数据的随机机制。推断是根据某一随机机制所产生的数据,分析判断它是否具有某些指定的属性,或者由多个不同的时间序列,分析不同的随机机制是否具有相同的属性。为此,我们必须假设可用适当的一类模型来描述产生观测数据的随机机制。在第三章中我们将介绍这类模型。预报是利用时间序列的相关性,预报随机机制在未来时刻的取值,本书将在第五章介绍时间序列的预报方法。控制是对某一随机机制(或多个随机机制)的一段观测数据的分析,寻求对某些量的控制,以达到优化目的。

## § 1.2 随机过程的基本概念

### 一、随机过程的定义

分析一个时间序列首先应对所给的数据选择合适的数学模型(或模型的类型)。如果将观测值  $x_t$  视为某个随机变量的样本值,那么时间序列  $\{x_t, t \in T\}$  就是一族随机变量  $\{X_t, t \in T\}$  的一个现实。对所给的数据用相应的数学模型近似描述,实质上是对某个随机过程的一个现实用相应的数学模型近似描述。下面给出随机过程的定义。

**定义 1.2.1** 设给定概率空间  $(\Omega, \mathcal{F}, P)$  及指标集  $T \subset (-\infty, \infty)$ , 如果对每一  $t \in T$ , 有一定义在  $(\Omega, \mathcal{F}, P)$  的随机变量  $X_t(\omega)$ ,  $\omega \in \Omega$  与之对应, 那么随机变量族

$$\{X_t(\omega), t \in T\} \quad (1.2.1)$$

称为定义在  $(\Omega, \mathcal{F}, P)$  上的随机过程。

常把(1.2.1)式简记为  $\{X_t, t \in T\}$  或  $\{X_t\}$ 。

注1 若指标集是整数集或非负整数集,则称 $\{X_t, t \in T\}$ 为随机序列,也称 $\{X_t, t \in T\}$ 为时间序列。

注2 由随机过程的定义可知,对每一固定的 $t \in T, X_t(\cdot)$ 是定义在 $(\Omega, \mathcal{F}, P)$ 上的随机变量;对每一固定的 $\omega \in \Omega, X_t(\omega)$ 是定义在 $T$ 上的实值函数,称此函数为随机过程 $\{X_t, t \in T\}$ 对应于此 $\omega$ 的一个样本函数(现实或轨道)。

注3 用时间序列这一术语既表示随机序列本身,也表示该过程的一个现实。

例1.2.1 设 $\{X_t, t = 1, 2, \dots\}$ 是取值为1或-1的独立随机变量序列,有

$$P\{X_t = 1\} = P\{X_t = -1\} = \frac{1}{2} \quad (1.2.2)$$

则存在概率空间 $(\Omega, \mathcal{F}, P)$ 及定义在 $\Omega$ 上的随机变量 $X_1, X_2, \dots$ ,使 $(X_1, \dots, X_n)$ 的联合概率分布

$$P\{X_1 = i_1, X_2 = i_2, \dots, X_n = i_n\} = 2^{-n} \quad (1.2.3)$$

其中 $i_j = 1$ 或 $i_j = -1, j = 1, \dots, n$ 。该随机过程的存在性可由Kolmogorov定理(§1.5)得证。

重复掷一硬币并记掷得正面为1,掷得反面为-1,所得到的时间序列可做为由(1.2.3)式定义的随机过程的一个现实。该过程的每一现实都是一系列1与-1。

例1.2.2 (随机相位和振幅的正弦波) 设 $A, \theta$ 是相互独立的随机变量,且 $A \geq 0, \theta$ 服从于 $[0, 2\pi]$ 上的均匀分布,即 $\theta \sim U[0, 2\pi]$ 。对任意给定的 $V \geq 0, r > 0$ ,随机过程 $\{X(t), t \in R\}$ 可由 $A$ 和 $\theta$ 来定义

$$X_t(\omega) = r^{-1}A(\omega)\cos(Vt + \theta(\omega)) \quad (1.2.4)$$

其中 $\omega \in \Omega$ 。

例1.2.3 (随机游动) 简单对称随机游动 $\{S_t, t = 0, 1, 2, \dots\}$

$$S_t = \sum_{i=1}^t X_i, t \geq 1 \quad (1.2.5)$$

其中 $S_0 = 0, X_i, i = 1, \dots, t$ 由例1.2.1定义。一般随机游动也可用(1.2.5)式定义,但 $X_1, X_2, \dots$ 是独立同分布随机变量。这类随机过程的存在性也可由Kolmogorov定理得证(习题1.10)。

例1.2.4 (分支过程) 分支过程是近似描述动物增长数据的数学模型。考虑某一群体,用 $X_i$ 表示第 $i$ 代中个体的数目。假设某一代的每一个个体可产生下一代的个数是独立同分布的随机变量,有 $X_0 = x$ (第0代人口总数),且

$$X_{t+1} = \sum_{j=1}^{x_t} Z_{t,j}, \quad t = 0, 1, 2, \dots \quad (1.2.6)$$

其中 $Z_{t,j}, t = 0, 1, 2, \dots, j = 1, 2, \dots$ 是独立同分布取非负整数值随机变量, $Z_{t,j}$ 表示第 $t$ 代第 $j$ 个个体产生的下一代的个数。

## 二、随机过程的有限维分布函数族和有限维特征函数族

设 $\{X_t, t \in T\}$ 是一随机过程,为了描述它的统计特征,需引入随机过程的有限维分布函数族的概念。

定义1.2.2 (随机过程 $\{X_t, t \in T\}$ 的有限维分布族) 令

$$J \triangleq \{t = (t_1, \dots, t_n); t = (t_1, \dots, t_n) \in T^n, t_1 < t_2 < \dots < t_n, n = 1, 2, \dots\},$$

称 $\{F_t(\cdot), t \in J\}$ 为随机过程 $\{X_t, t \in T\}$ 的有限分布函数族,特别当 $t = (t_1, \dots, t_n)$ 时

$$F_t(x) = P\{X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n\}, x = (x_1, \dots, x_n)^T \in R^n \quad (1.2.7)$$

为随机过程  $\{X_t, t \in T\}$  的  $n$  维分布函数族。

**定理 1.2.1** (Kolmogorov 定理) 分布函数族  $\{F_t(\cdot), t \in J\}$  是随机过程  $\{X_t, t \in T\}$  的有限维分布函数族的充分必要条件是对任意  $n \in \{1, 2, \dots\}$ ,  $t = (t_1, \dots, t_n) \in J$  以及  $1 \leq i \leq n$ , 有

$$\lim_{t_i \rightarrow \infty} F_t(x) = F_{t(i)}(x(i)) \quad (1.2.8)$$

成立, 其中  $t(i)$  和  $x(i)$  是分别由  $t$  和  $x$  删去第  $i$  个元素而得到的  $(n-1)$  维向量。

定理的证明见参考文献[1]。

令  $\varphi_t(\cdot)$  是相应于  $F_t(\cdot)$  的特征函数, 则称  $\{\varphi_t(\cdot), t \in J\}$  为随机过程  $\{X_t, t \in T\}$  的有限维特征函数族。特别当  $t = (t_1, \dots, t_n)$  时, 称

$$\varphi_t(u) = \int_{R^n} e^{iu^T x} dF_t(x_1, \dots, x_n), u = (u_1, \dots, u_n)^T \in R^n$$

为随机过程  $\{X_t, t \in T\}$  的  $n$  维特征函数。(1.2.8) 式与下式等价:

$$\lim_{u_i \rightarrow 0} \varphi_t(u) = \varphi_{t(i)}(u(i)) \quad (1.2.9)$$

其中  $u(i)$  是由  $u$  删去第  $i$  个元素得到的  $(n-1)$  维向量。

条件(1.2.8) 式称为有限维分布函数族的相容性。

### 三、宽平稳过程与严平稳过程

与时间序列分析的理论和方法有密切关系是平稳随机过程的概念。首先介绍随机过程的数字特征。设  $\{X_t, t \in T\}$  为随机过程, 则  $\{X_t\}$  的均值函数  $m_x(\cdot)$  定义为

$$m_x(t) = EX_t, \quad t \in T$$

其自协方差函数  $\gamma_x(\cdot, \cdot)$  定义为

$$\gamma_x(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - m_x(r))(X_s - m_x(s))], r, s \in T \quad (1.2.10)$$

称  $D_x(t) = \gamma_x(t, t)$ ,  $t \in T$  为  $\{X_t\}$  的方差函数。

如果  $\{X_t\}$  满足

$$X_t = Y_t + iZ_t, t \in T$$

其中  $\{Y_t, t \in T\}, \{Z_t, t \in T\}$  为实值随机过程, 则称  $\{X_t, t \in T\}$  为复值随机过程, 此时(1.2.10) 式为

$$\gamma_x(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - m_x(r))(\overline{X_s - m_x(s)})], r, s \in T \quad (1.2.11)$$

$$D_x(t) = E|X_t - m_x(t)|^2, t \in T,$$

其中符号  $\bar{\alpha}$  表示  $\alpha$  的共轭复数。由 Cauchy - Schwarz 不等式知, 如果  $E|X_t|^2 < \infty, t \in T$ , 则  $m_x(\cdot)$  和  $\gamma_x(\cdot, \cdot)$  都存在。

**定义 1.2.3** 设  $\{X_t, t \in T\}$  是复值随机过程, 其均值函数与协方差函数存在, 并满足

(1)  $EX_t = m$  (常数), 对一切  $t \in T$ 。

(2)  $\gamma_x(r, s) = \gamma_x(r+t, s+t)$ , 对一切  $r \in T, s \in T, r+t \in T, s+t \in T$  成立, 则称  $\{X_t, t \in T\}$  为宽平稳过程。

**注 1** 如果指标集  $T = \mathbf{Z} = \{0, \pm 1, \pm 2, \dots\}$ , 则称满足定义 1.2.3 条件的  $\{X_t, t \in \mathbf{Z}\}$  为宽平稳时间序列。

注2 设  $\{X_t, t \in T\}$  是宽平稳过程, 则有

$$\gamma_x(r, s) = \gamma_x(r - s, 0), \quad r, s \in T$$

$$\text{Cov}(X_{t+h}, X_t) = \gamma_x(h, 0) = \gamma_x(h), \quad t, t+h \in T$$

这表明宽平稳过程的自协方差函数是两个时刻间的间隔的一元函数。宽平稳过程定义中的(1)表明过程的均值不随时间的推移而变化; (2)表明过程前后两个时刻的统计意义下的线性相关性的强弱只依赖这两个时刻间的间隔而与它所处的位置  $t$  无关。在工程意义上, 宽平稳过程反映了一个系统处于稳态工作条件下的统计性质(稳态并不是指固定不变, 系统的工作仍具随机性)。

注3 称

$$\rho_x(h) = \frac{\gamma_x(h)}{\gamma_x(0)} = \text{Corr}(X_{t+h}, X_t), \quad t, t+h \in T \quad (1.2.12)$$

为宽平稳过程的自相关系数函数。

注4 称

$$R_x(h) = EX_{t+h}\bar{X}_t \quad (1.2.13)$$

为宽平稳过程的自相关函数。显然有  $R_x(h) = \gamma_x(h) + |m|^2$  成立。当  $m = 0$  时, 自相关函数与自协方差函数等价。

定义 1.2.4 设  $\{X_t, t \in T\}$  是随机过程, 如果对任意正整数  $n, t_1, t_2, \dots, t_n \in T$  及  $h(t_i + h) \in T, i = 1, 2, \dots, n$ ,  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  和  $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$  具有相同的分布函数, 则称  $\{X_t, t \in T\}$  为严平稳过程。

比较定义 1.2.3 和定义 1.2.4 可以看出, 宽平稳过程对时间推移的不变性(不依赖于  $t$ ) 表现在统计平均的一、二阶矩上, 而严平稳过程对时间推移的不变性则表现在概率分布上, 显然, 这两者的要求是不一样的, 应注意以下三点:

(1) 二阶矩有限的严平稳过程是宽平稳过程。因为

$$EX_t = \int_R x dF_t(x) = \int_R x dF_{t+x}(x) = m$$

$$EX_{t+h}\bar{X}_t = \iint_{R^2} x_1 \bar{x}_2 dF_{t+h,t}(x_1, x_2) = \iint_{R^2} x_1 \bar{x}_2 dF_{h,0}(x_1, x_2) = R_x(h)$$

与  $t$  无关。

(2) 宽平稳过程未必是严平稳过程, 例如, 设  $\{X_t, t \in \mathbb{Z}^+\}$  是独立随机变量序列, 其中  $\mathbb{Z}^+ = \{1, 2, 3, \dots\}$ , 当  $t$  是奇数时,  $X_t$  服从于均值为 1 的指数分布; 当  $t$  是偶数时,  $X_t \sim N(1, 1)$ , 则  $\{X_t, t \in \mathbb{Z}^+\}$  是宽平稳过程, 且

$$\gamma_x(h) = \begin{cases} 1, & h = 0 \\ 0, & h \neq 0 \end{cases}$$

但  $X_1$  与  $X_2$  的分布函数不相等。故  $\{X_t, t \in \mathbb{Z}^+\}$  不是严平稳过程。

(3) 严平稳过程未必是宽平稳过程。因为定义 1.2.4 中并不断言二阶矩存在, 而定义 1.2.3 的前提是协方差函数存在。

本书主要研究对象是宽平稳时间序列, 简称为平稳时间序列, 如不声明均假定平稳序列的均值是零。

例 1.2.1 (白噪声序列) 设  $\{Z_t, t \in \mathbb{Z}\}$  为互不相关的实值随机变量序列,  $EZ_t = 0$ , 其自协

方差函数为

$$\gamma(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \neq 0 \end{cases} \quad (1.2.14)$$

则称  $\{X_t, t \in \mathbb{Z}\}$  是均值为零, 方差为  $\sigma^2$  的白噪声序列, 简记为  $\{Z_t\} \sim WN(0, \sigma^2)$ 。显然白噪声序列是平稳时间序列。

如果  $\{Z_t, t \in T\}$  是均值为零、方差为  $\sigma^2$  的相互独立同正态分布的随机变量序列, 则记  $\{Z_t\} \sim IIDN(0, \sigma^2)$ 。

例 1.2.2 (滑动平均序列) 设  $\{Z_t\} \sim WN(0, \sigma^2)$ , 定义

$$X_t = \sum_{k=0}^q \theta_k Z_{t-k}, t = 0, \pm 1, \dots \quad (1.2.15)$$

其中  $\{\theta_k\}$  是实数列,  $k$  为一正整数, 则  $\{X_t, t \in \mathbb{Z}\}$  是平稳时间序列。事实上

$$\begin{aligned} EX_t &= \sum_{k=0}^q \theta_k E(Z_{t-k}) = 0 \\ \text{Cov}(X_{t+h}, X_t) &= E(X_{t+h}X_t) = \sum_{k=0}^q \sum_{j=0}^q \theta_k \theta_j E(Z_{t+h-k}Z_{t-j}) \\ &= \begin{cases} \sigma^2 \sum_{j=\max(0, -h)}^{\min(q, q-h)} \theta_j \theta_{j+h}, & 0 \leq |h| \leq q \\ 0, & |h| > q \end{cases} \end{aligned}$$

与  $h$  无关, 称  $\{X_t, t \in \mathbb{Z}\}$  为  $q$  阶滑动平均序列。在工程技术中设计如下一个滤波器, 它的滤波系数(实数)为  $\theta_0, \theta_1, \dots, \theta_q$ , 将白噪声序列  $\{Z_t, t \in \mathbb{Z}\}$  作为该滤波器的输入, 则其输出为滑动平均序列  $\{X_t, t \in \mathbb{Z}\}$ 。

例 1.2.3 设  $X_t = A\cos(\theta t) + B\sin(\theta t)$ , 其中  $A, B$  是互不相关的随机变量,  $EA = EB = 0, D(A) = D(B) = 1, \theta \in [-\pi, \pi]$ , 则随机过程  $\{X_t, t \in T\}$  是平稳过程。事实上

$$\begin{aligned} EX_t &= 0 \\ \text{Cov}(X_{t+h}, X_t) &= \text{Cov}(A\cos(\theta(t+h)) + B\sin(\theta(t+h)), A\cos(\theta t) + B\sin(\theta t)) \\ &= \cos(\theta t)\cos(\theta(t+h)) + \sin(\theta t)\sin(\theta(t+h)) \\ &= \cos(\theta h) \end{aligned}$$

与  $t$  无关。

例 1.2.4 设  $\{X_t, t \in \mathbb{Z}^+\}$  是随机游动,  $\{S_t = \sum_{i=1}^t X_i, t \geq 1\}$ , 其中  $X_1, X_2, \dots$  是独立同分布随机变量,  $EX_i = 0, DX_i = \sigma^2, i = 1, 2, \dots$ , 当  $h > 0$  时,

$$\text{Cov}(S_{t+h}, S_t) = \text{Cov}\left(\sum_{i=1}^{t+h} X_i, \sum_{j=1}^t X_j\right) = \sigma^2 t$$

故  $\{S_t, t \in \mathbb{Z}^+\}$  不是平稳时间序列。

#### 四、正态过程

定义 1.2.5 设  $\{X_t, t \in T\}$  是实随机过程, 如果对任意正整数  $n$  及  $t_1, t_2, \dots, t_n \in T$ ,  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  是  $n$  维正态随机向量, 则称  $\{X_t\}$  为正态过程。

正态过程有如下等价定义: