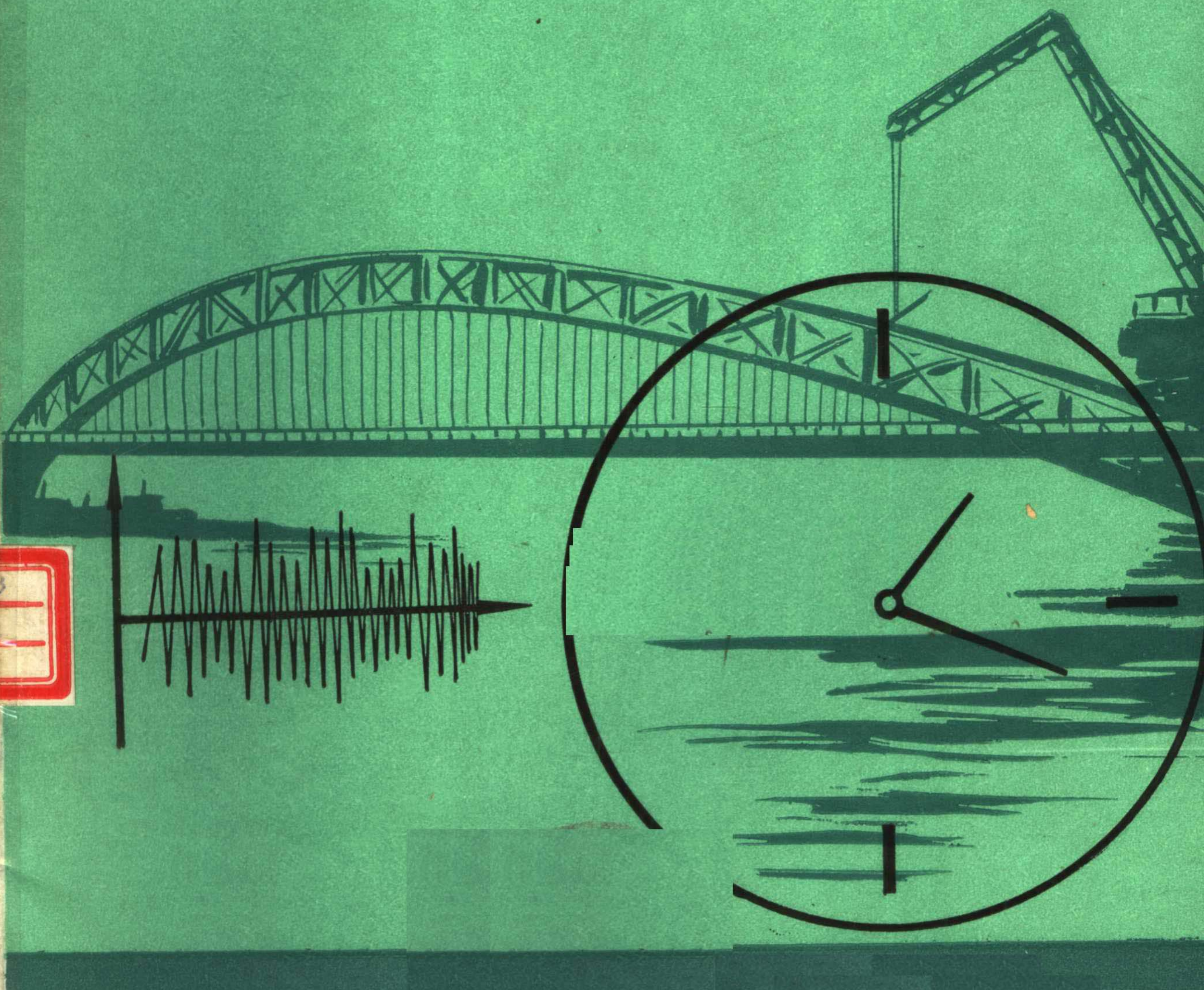
时序分析与工程应用

佟德纯　张黛华　编著



上海交通大学出版社

时序分析与工程应用

佟德纯　张黛华　编著

上海交通大学出版社

内 容 简 介

本书介绍了时间序列分析的基本原理、方法和应用，是一本说理简明、应用实例丰富的书。全书共分九章：第一章介绍了时序分析的数学基础；第二章至第六章论述了时序建模与参数估计的理论和方法；第七、八章讨论了时间序列的预报与谱估计；第九章重点分析了时序分析在设备故障诊断、模态参数识别、系统状态分类、预测预报以及生物信号处理等方面的应用实例。

本书可作为从事动态数据处理的机械、电力、石化、地震、宇航、生物、医学、经济、工业企业管理等方面工程技术人员的参考书，亦可作为高等院校相应专业高年级学生和研究生的教材。

时序分析与工程应用

出 版：上海交通大学出版社
（淮海中路1984弄19号）
印 刷：江苏省无锡县第二印刷厂
开 本：787×1092（毫米）1/16
印 张： 11.875
字 数： 304000
版 次：1990年 3 月 第 1 版
印 次：1990年 3 月 第 1 次
印 数：1—2000
ISBN7—313—00684—5/TP·29
成本定价：6.50元

序　言

随着电子计算机技术的发展，动态数据处理和实测的信号分析已形成了一门新兴的学科，它广泛应用于各种科学和技术领域中，根据数学模型的特点，分为以FFT分析技术为基础的非参数模型法和以时间序列建模的参数模型法。前者称为“工程信号处理”，后者称为“时间序列分析”，简称为“时序分析”。

有序的随机数据称为时间序列，时序分析实际上是对有序的随机数据进行分析、研究与处理的一种方法。这种观测到的数据就是依时间顺序排列，並各具大小。正是这种有序性和大小表达了数据中所包含的信息，反映了数据内部的相互联系和规律性，蕴含了产生这些数据的现象、过程，以及系统的有关特性。因此，分析、研究与处理这些有序的随机数据，正是为了揭示和提取它所包含的信息，掌握它内部的规律，了解有关系统的特性，进而推断和预测系统的未来变化。

时序分析是应用数学的一个重要分支，是概率统计学的重要组成部分。近十多年来，在理论和应用上都迅速发展，尤其是在工程应用方面更为突出。它所提供的动态数据处理方法，在国内外的应用正日趋广泛，随着电子计算机的普及，时序分析方法越来越显示出它的重要性和实用性，从而越来越为广大科学工作者和工程技术人员所关注。

在自然科学、工程技术、以及社会科学的许多领域里，实际工作者和科研人员常常要对一系列的观测数据进行分析研究，因此，对时序分析方面的书籍的需要是显而易见的。我们多年来从事这方面的教学和科研工作，深知这种需要的迫切性，所以编写了《时序分析与工程应用》一书。编写过程中，在介绍时间序列分析的基本原理和方法时着重于具体应用和实例分析。

全书共分九章，内容大体可分为四大部分：1.时序分析的数学基础；2.时序建模与参数估计；3.时间序列的预报与谱估计；4.工程应用实例分析。

其中第三、七、八、九章由佟德纯编著，第一、二、四、五、六章由张黛华编著，全书是在相互校审基础上定稿的。我们希望这本书能在应用时序分析方法解决实际工程问题上，对读者有所帮助。

限于作者水平，书中定有许多不妥之处，竭诚欢迎读者批评指正。

编著者　1989年8月

目　录

第一章　动态数据建模基础知识

第一节　向量和矩阵

一、向量及其基本运算

将n个数a_1，$a_2\cdots a_n$按一定次序作如下排列

$$a_1,\ a_2,\ \cdots a_n$$

称为n维向量或n元向量。数$a_i(i=1,\ 2,\ \cdots n)$叫做n维向量的第i个分量或坐标，本书用相应字母的黑体表示向量，如

$$\mathbf{a}=(a_1,\ a_2\cdots a_n)$$

分量都是0的向量，即（0，0，…0）叫做零向量。

下面定义n维向量的运算

n维向量的运算可由三维向量的运算推广而来。如令

$\mathbf{a}=(a_1,\ a_2,\ \cdots a_n)$，$\mathbf{b}=(b_1,\ b_2\cdots b_n)$，$\mathbf{c}=(c_1,\ c_2,\ \cdots c_n)$

于是有如下结论

1. 如果两个向量$\mathbf{a}$和$\mathbf{b}$的对应分量相等，则称$\mathbf{a}$和$\mathbf{b}$相等，记为

$$\mathbf{a}=\mathbf{b}$$

2. 两个向量$\mathbf{a}$和$\mathbf{b}$的加法，是把它们的对应分量相加，所得的向量称为$\mathbf{a}$与$\mathbf{b}$的和。记为$\mathbf{a}+\mathbf{b}$，即

$$\mathbf{a}+\mathbf{b}=(a_1+b_1,\ a_2+b_2,\ \cdots a_n+b_n)$$

3. 数k与向量$\mathbf{a}$的乘法，是用k乘它的每一个分量，所得的向量称为数k与$\mathbf{a}$的乘积，记为$k\mathbf{a}$，即

$$k\mathbf{a}=(ka_1,\ ka_2,\ \cdots ka_n)$$

4. 两个向量$\mathbf{a}$与$\mathbf{b}$的减法，定义为$\mathbf{a}+(-1)\mathbf{b}$。记为$\mathbf{a}-\mathbf{b}$，即

$$\mathbf{a}-\mathbf{b}=(a_1-b_1,\ a-b_2,\ \cdots a_n-b_n)$$

对于n维向量的加、减、数乘与乘法的运算，显然满足下面规律

(1)　$\mathbf{a}+\mathbf{b}+\mathbf{c}=(\mathbf{a}+\mathbf{b})+\mathbf{c}$；

(2)　$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$；

(3)　$\mathbf{o}+\mathbf{a}=\mathbf{a}+\mathbf{o}=\mathbf{a}$（设$\mathbf{o}$为零向量，下同）；

(4)　$\mathbf{a}+(-\mathbf{a})=(-\mathbf{a})+(\mathbf{a})=\mathbf{o}$；

(5)　$k(\mathbf{a}+\mathbf{b})=k\mathbf{a}+k\mathbf{b}$；

(6)　$(k+l)\mathbf{a}=k\mathbf{a}+l\mathbf{a}$（$l$为任意数）；

(7)　$(kl)\mathbf{a}=k(l\mathbf{a})$；

(8)　$1\mathbf{a}=\mathbf{a}$。

5. 把向量$\mathbf{a}$和$\mathbf{b}$各分量的乘积$a_ib_i(i=1,\ 2,\ \cdots n)$加起来得到一个数，称此数为$\mathbf{a}$与$\mathbf{b}$的内积。记为$(\mathbf{a},\ \mathbf{b})$，即

$$(a,\ b)=a_1b_1+a_2b_2+\cdots a_nb_n=\sum_{i=1}^{n}a_ib_i$$

如果$(a,\ b)=0$，则称a与b是互相正交的。显然有$(0,\ a)=0$，即零向量与任一向量正交。称数值$\sqrt{(a,\ a)}$为向量a的模或长度。记为$|a|$，即

$$|a|=\sqrt{(a,\ a)}=\sqrt{a_1{}^2a_2{}^2+\cdots+a_n{}^2}=\sqrt{\sum_{i=1}^{n}a_i{}^2}$$

模是1的向量叫做单位向量，因而向量$\frac{a}{|a|}$就是模为1的、方向与a相一致的单位向量。

由两个向量内积的定义，容易推得内积具有下述性质

(1)　$(a+b,\ c)=(a,\ c)+(b,\ c)$；

(2)　$(a,\ b)=(b,\ a)$；

(3)　$(ka,\ b)=k(a,\ b)$；

(4)　$(a,\ a)\geqslant 0$，当$a=0$时，等号成立。

二、矩阵的概念及其基本运算

取$n\times m$个实数a_{ij}，$i=1,\ 2\cdots n$；$j=1,\ 2,\ \cdots m$，作如下排列，称为一个n行m列的矩阵，简称为$n\times m$矩阵。即

$$\begin{pmatrix} a_{11} & a_{12}\cdots a_{1m} \\ a_{21} & a_{22}\cdots a_{2m} \\ a_{n1} & a_{n2}\cdots a_{nm} \end{pmatrix}$$

矩阵常用大写的拉丁字母A，B，C，…表示，于是上述$n\times m$矩阵可简记为

$$A=\begin{pmatrix} a_{11} & a_{12}\cdots & a_{1m} \\ a_{21} & a_{22}\cdots & a_{2m} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2}\cdots & a_{nm} \end{pmatrix}(a_{ij})_{n\times m}=(a_{ij})$$

当$n=m$时，称它为n阶方阵，a_{ij}称为A的第i行第j列元素，或$(i,\ j)$元。各元素全为0的$n\times m$矩阵简记为0，称它为零矩阵。当$m=1$时，A成为

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$$

称这样的矩阵为列矩阵或列向量。同样，当$n=1$时，便得

$$(a_{11},\ a_{12},\ \cdots a_{1m})$$

称这样的矩阵为行矩阵或行向量。

将矩阵$A=(a_{ij})_{n\times m}$的行与列互相调换所成的矩阵称为A的转置矩阵。记为A^T

$$A^T=\begin{pmatrix} a_{11} & a_{12}\cdots & a_{1m} \\ a_{21} & a_{22}\cdots & a_{2m} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2}\cdots & a_{nm} \end{pmatrix}^T=\begin{pmatrix} a_{11} & a_{21}\cdots & a_{n1} \\ a_{12} & a_{22}\cdots & a_{n2} \\ \vdots & \vdots & \vdots \\ a_{1m} & a_{2m}\cdots & a_{nm} \end{pmatrix}$$

如果A是$m\times n$矩阵，那末A^T是$n\times m$矩阵。

矩阵的转置运算满足下列性质

1. $(A+B)^T=A^T+B^T$;

2. $(A^T)^T=A$;

3. $(kA)^T=kA^T$(k为任一实数)；

4. $(AB)^T=B^TA^T$。

下面定义矩阵的运算

1. 矩阵的加法

两个$n\times m$矩阵$A=(a_{ij})$，$B=(b_{ij})$的和，记作$A+B$，仍是$n\times m$矩阵。即

$$A+B=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12}\cdots a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22}\cdots a_{2m}+b_{2m} \\ \vdots & \\ a_{n1}+b_{n1} & a_{n2}+b_{n2}\cdots a_{nm}+b_{nm} \end{pmatrix}$$

容易验证，矩阵加法具有以下性质

(1) $A+B=B+A$;

(2) $(A+B)+C=A+(B+C)$;

(3) $C(A+B)=CA+CB$,

$(A+B)C=AC+BC$;

(4) $A+0=0+A=A$。

2. 数乘和矩阵的乘法

任一实数k和$n\times m$矩阵A的乘积记为kA，仍是$n\times m$矩阵。即

$$kA=\begin{pmatrix} ka_{11} & ka_{12}\cdots ka_{1m} \\ ka_{21} & ka_{22}\cdots ka_{2m} \\ \vdots & \vdots \quad\quad \vdots \\ ka_{n1} & ka_{n2}\cdots ka_{nm} \end{pmatrix}$$

$n\times m$矩阵$A=(a_{ij})$和$m\times l$矩阵$B=(b_{jk})$的乘积记为AB，是$n\times l$矩阵，它的第(i, k)元素为$\sum_{j=1}^{m}a_{ij}b_{jk}$，即

$$AB=\sum_{j=1}^{m}a_{ij}\ b_{jk}$$

容易验证，矩阵的乘法具有以下性质

(1) $1A=A$;

(2) $(kl)A=k(lA)$，$(kA)B=k(AB)$，(其中k，l为常数，下同)。

(3) $k(A+B)=kA+kB$;

(4) $(AB)C=A(BC)$;

(5) $(k+l)A=kA+lA$。

注意，矩阵乘法不满足交换律，即$AB=BA$一般不成立。

3. 矩阵的减法

两个$n\times m$矩阵$A=(a_{ij})$与$B=(b_{ij})$之差，定义为$A+(-1)B$，记为$A-B$。即

$$A-B=A+(-1)B=\begin{pmatrix} a_{11}-b_{11} & a_{12}-a_{12} & \cdots a_{1m}-b_{1m} \\ a_{21}-b_{21} & a_{22}-a_{22} & \cdots a_{2m}-b_{2m} \\ \vdots & \vdots & \vdots \\ a_{n1}-b_{n1} & a_{n2}-b_{n2} & \cdots a_{nm}-b_{nm} \end{pmatrix}$$

在矩阵运算中，对于较复杂的矩阵，在计算过程中采用"矩阵分块"法，它可以将一些高阶矩阵的运算简化为n个较低阶矩阵的运算。例如，将两个$n\times m$矩阵A、B分别分成四块

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \qquad B=\begin{pmatrix} B_{11} & B_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

其中A_{ij}，B_{ij}为$n_i\times m_j$矩阵，$i=1,\ 2$；$j=1,\ 2$，$n_1+n_2=n$；$m_1+m_2=m$，则

$$A+B=\begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22} \end{pmatrix}$$

$$kA=\begin{pmatrix} kA_{11} & kA_{12} \\ kA_{21} & kA_{22} \end{pmatrix},\quad A^T=\begin{pmatrix} A^T{}_{11} & A^T{}_{12} \\ A^T{}_{21} & A^T{}_{22} \end{pmatrix}$$

对于矩阵乘法，仍然成立。例如，将$n\times m$矩阵A仍分成如上的四块，而将$m\times l$矩阵C分成如下四块

$$C=\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

其中C_{ij}为$m_i\times L_j$；$i=1,\ 2$；$j=1,\ 2$；$L_1+L_2=L$，则有

$$AC=\begin{pmatrix} A_{11}C_{11}+A_{12}C_{21} & A_{11}C_{12}+A_{12}C_{22} \\ A_{21}C_{11}+A_{22}C_{21} & A_{21}C_{12}+A_{22}C_{22} \end{pmatrix}$$

显然，一个给定的矩阵可用不同的方法分块，行块和列块的数目也不必相等。

三、方阵

行数等于列数的矩阵称为方阵，n行n列的方阵称为n阶方阵。主对角线元全为1，其余元全为0的n阶方阵称为单位阵。记为I_n或I，即

$$I_n=\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}_{n\times n}$$

它同任何一个n阶方阵$A=(a_{ij})$相乘有

$$I_nA=A=AI_n$$

设$|A|$为n阶方阵A的行列式，B也是n阶方阵，则有

$|A^T|=|A|$，$|kA|=k^n|A|$（k为任一实数），$|AB|=|A|\ |B|$。

设A为n阶方阵，如果由A的元所构成的行列式$|A|\neq0$，称A为非奇异方阵，否则称为奇异方阵。

如果方阵A是非奇异的，那末A有逆阵A^{-1}，并且

$$A^{-1}=(-1)^{i+j}\frac{|A_{ji}|}{|A|}$$

如果A^{-1}是A的逆阵，又B也是A的逆阵，则有

$$A^{-1}A=AA^{-1}=I_n,\ \ BA=AB=I_n$$

那末

$$B=BI_n=B(AA^{-1})=(BA)A^{-1}=I_nA^{-1}=A^{-1}$$

下面给出关于逆阵运算的若干结果

1. $|A^{-1}|=|A|^{-1}$；
2. $(A^T)^{-1}=(A^{-1})^T$；
3. $(A^{-1})^{-1}=A$；
4. $(AB)^{-1}=B^{-1}A^{-1}$；
5. $A^{-k}=(A^{-1})^k$，$(k>0)$；
6. $(kA)^{-1}=k^{-1}A^{-1}$（k为任一实数）。

n阶方阵主对角线元之和称为n阶方阵的迹，记为T_rA。即

$$TrA=\sum_{i=1}^{n}a_{ii}$$

于是有

$$TrA^T=TrA;$$

$$Tr(A+B)=TrA+TrB;$$

$$Tr(kA)=kTrA。$$

如果A，B分别是$n\times m$和$m\times n$矩阵，则有

$$Tr(AB)=\sum_{i=1}^{n}\left(\sum_{j=1}^{m}a_{ij}b_{ji}\right)=\sum_{j=1}^{m}\left(\sum_{i=1}^{n}b_{ji}a_{ij}\right)=Tr(BA)$$

又如果α是n维向量，A是n阶方阵，则有

$$\alpha^TA\alpha=\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_ia_{ij}\alpha_j=Tr(A\alpha\alpha^T)$$

其中

$$\alpha=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}$$

下面给出求分块方阵的行列式和逆的公式

设A是$k+l$阶分块矩阵，即

$$A=\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{pmatrix}或A=\begin{pmatrix}A_{11} & 0\\ A_{21} & A_{22}\end{pmatrix}$$

其中A_{11}和A_{22}分别为k阶和l阶方阵，A_{12}和A_{21}分别为$k\times l$阶和$l\times k$阶矩阵。由行列式性质知，$|A|=|A_{11}|\,|A_{22}|$，在A_{11}和A_{22}都是可逆阵条件下，可以得到

$$\begin{aligned}|A|&=|A_{11}|\,|A_{22}-A_{21}A_{11}{}^{-1}A_{12}|\\&=|A_{11}-A_{12}A_{22}{}^{-1}A_{21}|\,|A_{22}|\end{aligned}\qquad(1-1)$$

根据三角阵求逆公式，A^{-1}可表示为

$$A^{-1}=\begin{pmatrix} A_{11}^{-1} & A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \\ 0 & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{pmatrix}\begin{pmatrix} I_k & 0 \\ -A_{21}A_{11}^{-1} & I_l \end{pmatrix}$$

$$=\begin{pmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}\widetilde{A}_{22}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\widetilde{A}_{22}^{-1} \\ -\widetilde{A}_{22}^{-1}A_{21}A_{11}^{-1} & \widetilde{A}_{22}^{-1} \end{pmatrix} \qquad (1-2)$$

或

$$A^{-1}=\begin{pmatrix} (A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & 0 \\ -A_{22}^{-1}A_{12}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{22}^{-1} \end{pmatrix}\begin{pmatrix} I_k & -A_{12}A_{22}^{-1} \\ 0 & I_l \end{pmatrix}$$

$$=\begin{pmatrix} \widetilde{A}_{11}^{-1} & -\widetilde{A}_{11}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}\widetilde{A}_{11}^{-1} & A_{22}^{-1}A_{21}\widetilde{A}_{11}^{-1}A_{12}A_{22}^{-1}+A_{22}^{-1} \end{pmatrix} \qquad (1-3)$$

式中 $\widetilde{A}_{22}=A_{22}-A_{21}A_{11}^{-1}A_{12}$

$\widetilde{A}_{11}=A_{11}-A_{12}A_{22}^{-1}A_{21}$

从而可得矩阵的反演公式（对比式(1－2)、式(1－3)的左上角块和右上角块得到）

$$(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}=A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{12}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} \qquad (1-4)$$

以及

$$A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}=(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \qquad (1-5)$$

对于公式（1－4）和（1－5），只要表示式中出现的逆都存在，总是成立。

四、矩阵的秩、范数与对称阵

1． 矩阵的秩

给定m个n维向量α_1，α_2，…α_m，如果存在不全为零的数，即

$$(k_1,\ k_2,\ \cdots k_m)\neq(0,\ 0,\ \cdots 0)$$

使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m=0$$

则称此向量组α_1，$\alpha_2\cdots\alpha_m$是线性相关；否则称为线性无关。

设α_{i1}，α_{i2}，α_{ir}为n维向量组α_1，α_2，…α_m中的一个线性无关部分组，如果任 意r个以上向量构成部分组均线性相关，则称α_{i1}，α_{i2}，…α_{ir}为α_1，α_2，…α_m的一个最大线性 无 关组。n维向量组的秩是这个向量组内的最大线性无关组的向量个数。矩阵的秩定义为它的全体行向量组(或列向量)的秩；亦可定义为矩阵所含不等于零的子行列式的最大阶数。n×m矩阵A的秩，记为rankA，且

$$\text{rank}A\leqslant\min(n,\ m)$$

如果等号成立，则称A为满秩矩阵。由于$|A|\neq0$的n阶方阵A为非奇异方阵，由矩阵秩的定义知，当$|A|\neq0$时，rankA就等于阶数n，因此，又称A为满秩方阵，它是可逆的。

2． 矩阵的范数

对于n×m矩阵A定义其范数为

$$||A||=(Tr(A^TA))^{\frac{1}{2}}=\left(\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^2\right)^{\frac{1}{2}} \qquad (1-6)$$

如果设α是n维向量，则

$$||\alpha||=(\alpha^T\alpha)^{\frac{1}{2}}=\left(\sum_{i=1}^{n}\alpha_i^2\right) \qquad (1-7)$$

范数有如下一些基本性质

(1) $||A|| = 0$当且仅当$A = 0$;

(2) $||A+B|| \leqslant ||A|| + ||B||$;

(3) $||kA|| = |k|\,||A||$（k为任一实数）;

(4) $||AB|| \leqslant ||A||\,||B||$。

3. 对称阵

n阶方阵A，如果满足$A^T = A$，则称A为对称阵。两个同阶对称阵之和、任一实数与对称阵之和以及可逆对称阵之和，仍是对称阵。

一个n阶对称阵A，如果对于任意n维向量α，成立

$$\alpha^T A \alpha = \sum_{i,j=1}^{n} a_{ij} x_i x_j \geqslant 0 \tag{1-8}$$

则称 A是非负定阵，记为 $A \geqslant 0$。如果对于任一 $\alpha \neq 0$，有 $\alpha^T A \alpha > 0$，则称A为正定阵，记为 $A > 0$。显然有

(1) 单位阵$I_n > 0$;

(2) 两个同阶非负定阵之和仍为非负定阵，若其中一个是正定阵，其和也是正定阵;

(3) 正实数与非负定（正定）阵之积仍然是非负定（正定）阵;

(4) 任何正定阵必为可逆阵，而且它的逆也是正定阵;

(5) 两个n阶对称阵A和B，如果满足$A-B \geqslant 0$，即对于任意n维向量α，如有$\alpha^T A \alpha \geqslant \alpha^T B \alpha$，则称A不小于B，记为$A \geqslant B$（或$B \leqslant A$）;

(6) 如果A是n阶非负定阵，C是$n \times m$矩阵，则易证$C^T AC$是m阶非负定阵，且$C^T AC$为对称阵。

五、矩阵微分运算

设$n \times m$矩阵A、B的元都是x的函数，$\lambda = \lambda(x)$，则矩阵加法、数乘和乘法的运算规则具有如下简单性质

1. $\frac{d}{dx}[A(x) \pm B(x)] = \frac{dA(x)}{dx} \pm \frac{dB(x)}{dx}$;

2. $\frac{d}{dx}[\lambda A(x)] = \lambda \frac{dA(x)}{dx}$;

3. $\frac{d}{dx}[A(x)B(x)] = A(x)\frac{dB(x)}{dx} + B(x)\frac{dA(x)}{dx}$

其中A、B对x的微商为如下的$n \times m$矩阵

$$\frac{dA}{dx} = \frac{da_{ij}}{dx}, \quad \frac{dB}{dx} = \frac{db_{ij}}{dx}$$

如设f、g是以$n \times m$矩阵X的nm个元α_{ij}为自变元的函数，则有

$$\frac{d}{d\alpha}(f+g) = \frac{df}{d\alpha} + \frac{dg}{d\alpha};$$

$$\frac{d}{d\alpha}(fg)=g\frac{df}{d\alpha}+f\frac{dg}{d\alpha}。$$

其中f（或g）对α的微商为如下的n×m矩阵

$$\frac{df}{d\alpha}=\frac{\partial f}{\partial \alpha_{ij}}$$

4. 设A为n阶方阵，其元是x的函数，则

$$\frac{dA^{-1}}{dx}=-A^{-1}\frac{dA}{dx}A^{-1};$$

$$\frac{d|A|}{dx}=|A|Tr(A^{-1}\frac{dA}{dx})。$$

5. 设α为n×m变元矩阵，A、B分别为n×n和n×m常数矩阵，则

$$\frac{dTr(B\alpha)}{d\alpha}=\frac{dTr(\alpha^T B^T)}{d\alpha}=B^T;$$

$$\frac{dTr(A\alpha)}{d\alpha}=(A+A^T)\alpha。$$

根据性质5，设α为n维向量（每一分量为变元），k为n维常数向量，A为n阶常数对称矩阵，则

$$\begin{aligned}\frac{d}{d\alpha}[(\alpha-k)^T A(\alpha-k)]&=\frac{d}{d\alpha}Tr[(\alpha-k)^T A(\alpha-k)]\\&=(A+A^T)(\alpha-k)\\&=2A(\alpha-k)。\end{aligned}$$

第二节　概率与统计估计基础

一、随机事件与概率

1. 随机事件

在某些不变的条件下重复进行试验或观测，可能发生也可能不发生的事件，称为随机事件，简称事件。在每次试验或观测中一定发生的事件称为必然事件，通常用Ω表示。在每次试验或观测中一定不发生的事件称为不可能事件，通常用φ表示。例如，投掷一枚分币，“正面朝上”这一事件，是随机事件；“上海地区8月份平均温度为2℃”这一事件，是不可能事件；“在标准大气压下，水加热到100℃时要沸腾”这一事件，是必然事件。

2. 概率的基本概念及其运算

设在某组不变的条件下，随机事件A在n次重复试验中发生m次，则称m/n为事件A在n次试验中发生的频率，显然$0\leqslant m/n\leqslant 1$。如果事件A是必然事件，则$m/n=1$；如果事件A是不可能事件，则$m/n=0$。由经验可知，当n足够大时，事件A发生的频率趋于某一稳定值，即频率徘徊于某一数值附近，这个数值称为事件A的概率，记为P(A)。概率论中最基本的有下面两条运算法则。

（1） 概率加法法则

加法定理 两互斥事件A、B之和的概率等于两事件A、B的概率之和。即若$AB=\varphi$，则有

$$P(A\cup B)=P(A)+P(B) \tag{1-9}$$

由加法定理推得如下一些结论

1) $P(A)=1-P(\bar{A})$； (1-10)

2) 若$A\subset B$，则$P(B-A)=P(B)-P(A)$；

3) 若A_1，A_2，…为两两互斥事件，则

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

4) 对于任意两件事件A、B则

$$P(A\cup B)=P(A)+P(B)-P(AB) \tag{1-11}$$

上式可推广到任意多个随机事件的情形。

(2) 概率乘法法则

乘法定理 两事件A、B积的概率等于其中一个事件的概率（概率不为零）与另一事件在前一事件已经发生条件下的条件概率的乘积。即

$$P(AB)=P(A)P(B|A) \quad P(A)\neq 0 \tag{1-12}$$

或

$$P(AB)=P(B)P(A|B) \quad P(B)\neq 0$$

显然，当A、B独立时有

$$P(AB)=P(A)P(B)$$

乘法定理不难推广到有限多个随机事件的情形。即

$$P(A_1A_2\cdots P_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})$$

下面叙述全概率公式和贝叶斯（Bayes）公式

设必然事件Ω是有限个或无限个两两互斥事件B_1，B_2，…的和，且构成完备事件组，$P(B_i)>0$，$i=1,2,\cdots$。则对任一事件A，当且仅当B_1，B_2…中的任一事件发生时A才发生，于是有

$$A=A\Omega=A(B_1+B_2+\cdots)=AB_1+AB_2+\cdots$$

由于各个B_i两两互斥，则由概率的加法定理和乘法定理可得如下关系式

$$\begin{aligned}P(A)&=P(AB_1)+P(AB_2)+\cdots\\&=\sum_{i=1}^{\infty}P(AB_i)\\&=\sum_{i=1}^{\infty}P(B_i)P(A|B_i)\end{aligned} \tag{1-13}$$

称(1-13)式为全概率公式。根据乘法定理有

$$P(A)P(B_i|A)=P(B_i)P(A|B_i)$$

由此得在任一事件A已发生的条件下，事件B_i发生的概率为

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{P(A)}=\frac{P(B_i)P(A|B_i)}{\sum_{i=1}^{\infty}P(B_i)P(A|B_i)} \tag{1-14}$$

称（1-14）式为贝叶斯（Bayes）公式

二、随机变量及其分布

用数学方法处理随机现象时，需要以数量来描述。设

$$X=\begin{cases}1 & \text{抽到的产品为正品；}\\ 0 & \text{抽到的产品为次品。}\end{cases}$$

如果在确定的生产条件下，已知这批产品的次品率为p，则

$$P\{X=0\}=p,\ P\{X=1\}=1-p$$

通常称这样的变量X为随机变量，称$P\{X=0\}=p$，$P\{X=1\}=1-p$为随机变量X的概率分布。所以当一个变量所取的值与随机试验的结果有关时便称为随机变量。

显然，可推广随机变量的可能取值为有穷个或可列无穷多个的情形。这时其概率分布为

$$P\{X=x_i\}=p_i\geqslant 0,\ i=1,\ 2,\ \cdots \qquad (1\text{-}15)$$

式中

$$\sum_{i=1}^{\infty} p_i=1$$

称这样的随机变量X为离散型随机变量。

随机变量的可能取值也可以是任何实数值。例如，测量误差、洪峰值、候车的等待时间、某个时刻的噪声值皆是，这类随机变量称为非离散型随机变量，非离散型随机变量中最常见的是连续型随机变量，研究连续型随机变量的分布时，通常是把随机变量的观测值所在的区间划分成若干个相邻的小区间，各小区间的长度可以相等，也可以不相等，这样就可得到随机变量落在各个小区间内的频率分布。为了研究随机变量的理论分布，引进分布函数的概念。

设x是任意实数，考虑随机变量X取小于x的值的概率，显然它是x的函数，记为

$$F(x)=P\{X\leqslant x\} \qquad (1\text{-}16)$$

称此函数为随机变量X的分布函数。

下面给出连续型随机变量的一般定义

设随机变量X的分布函数为F(x)，如果存在一个非负函数f(x)，使对任一实数x都有

$$F(x)=\int_{-\infty}^{x} f(t)dt \qquad (1\text{-}17)$$

则称此随机变量X为一个连续型随机变量，函数f(x)称为连续型随机变量X的概率密度函数（或概率密度）。

由概率分布的性质知

$$f(x)\geqslant 0,\qquad \int_{-\infty}^{\infty} f(x)dx=1$$

下面列举两种常用的概率密度函数

正态分布的概率密度函数（见图1-1）

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \qquad (1\text{-}18)$$

$$-\infty<x<\infty$$

显然f(x)的峰点在μ处，μ±σ处有两个拐点。

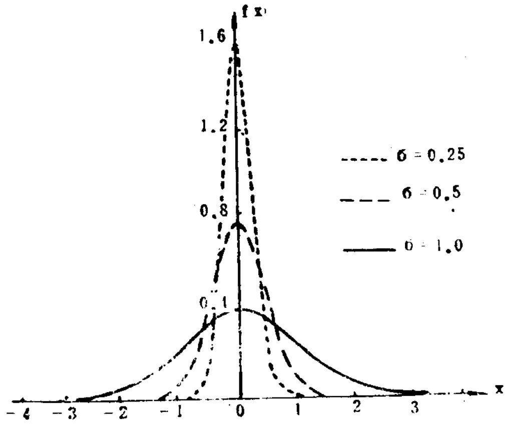


图 1－1　正态分布的概率密度函数

如果随机变量X具有上述概率密度函数，则称此随机变量X服从正态分布，常用符号$X \sim N(\mu, \sigma^2)$表示。当$\mu = 0$，$\sigma = 1$时，有

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \tag{1-19}$$

则称随机变量X服从标准正态分布，常用符号$X \sim N(0, 1)$表示。

均匀分布的概率密度函数（见图1-2）

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leqslant x \leqslant b \\ 0 & \text{其它} \end{cases} \tag{1-20}$$

如果随机变量X在固定区间[a, b]上具有上述概率密度函数，则称此随机变量X在区间[a,b]上服从均匀分布。

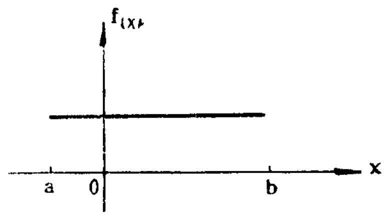


图 1－2　均匀分布的概率密度函数

三、随机向量及其分布

在生产实际与理论研究中，常常会遇到需要同时用几个随机变量才能较好地描绘某一试验（或随机现象）。例如在n批具有相同次品率p的产品中，各随机抽取一件产品，以$X = (X_1, X_2, \cdots X_n)$表示它们分别抽到为正品、次品的结果，这时X就是一个多维随机变量，简称随机向量，它可能取值的集合是其分量取值为1或0的n维向量全体。

类似于一维随机变量，称函数