普通高等教育测绘类规划教材

广义测量平差

（第二版）

崔希璋　於宗俦　陶本藻　刘大杰　于正林

测绘出版社

普通高等教育测绘类规划教材

广义测量平差

（第二版）

崔希璋　於宗俦　陶本藻

刘大杰　于正林

测绘出版社

（京）新登字 065 号

广义测量平差（第二版）

崔希璋　於宗俦　陶本藻

刘大杰　于正林

*

测绘出版社出版

北京大兴星海印刷厂印刷

新华书店总店科技发行所发行

*

开本 787×1092　1/16 · 印张 19.5 · 字数 440 千字

1982 年 6 月第一版 · 1992 年 10 月第二版 · 1992 年 10 月第四次印刷

印数 13 501—15 500册 · 定价 5.10 元

ISBN 7-5030-0521-1/P·196

前 言

随着近代统计估计理论、矩阵代数和电子计算机的发展，同时也由于物理大地测量、测绘空间技术和地球动力学等领域对数据处理的需要，测量平差的理论和方法有了很大的发展。在这种形势下，我们在1982年编写出版了《广义测量平差》一书。近十年来，测量平差的理论和方法又有了新的发展；为了提高测绘专业本科的教学，《广义测量平差》中的部分内容已经充实到本科教材《测量平差基础》(增订本)中。由于这些情况的改变，我们对《广义测量平差》一书的内容进行了一次修改和增删，以满足和适应当前研究生学习（包括助教进修生）和广大测绘科技工作者的需要。

这本《广义测量平差》(第二版)与1982年的版本相比，除了部分内容已经充实到本科教材而此书删去之外，有的章节作了修改，有的则是这次新增和补充的内容。

作了修改的章节有：第二章"估计方法和广义测量平差原理"和第五章"滤波与配置"；这次新增和补充的章节有：第一章"随机过程基础"，主要包括随机过程的基本概念，平稳随机过程，协方差函数与相关函数，以及几个特殊随机过程等内容；第三章"秩亏自由网平差"，着重讲了有关平差的基准问题，介绍了几种常用自由网的平差方法，此外还介绍了自由网的动态平差和分区平差，附有系统参数的自由网平差，以及自由网平差的特点等内容；第四章"平差随机模型的验后估计"，主要是介绍了现有的几种估计方法，如赫尔默特估计法和最小范数二次无偏估计法(即MINQUE法)，还给出了方差分量估计中的精度评定公式；第六章"若干平差问题的补充"，主要包括法方程系数阵中部分主子块秩亏时的解算方法，各种验后协因数阵的秩，顾及起始数据误差时的精度评定，以及先验协因数阵为奇异时的平差等个别问题。为了进一步加深理解平差问题的实质，本章还简要地讨论了在向量空间中的平差问题、广义G-M模型下的平差问题及等价观测值的概念等；第七章"动态线性系统的卡尔曼滤波"和第八章"卡尔曼滤波的渐近性质和误差分析"，对动态线性系统数学模型的建立、卡尔曼滤波方法及其应用等问题作了较为系统而详细的论述，同时也讨论了有关卡尔曼滤波的稳定性和误差分析等问题。在上述新增和补充的各章中，其内容除了取材于国外已有的文献之外，有的则是多年来我们在科研中所取得的部分成果。

本书可作为测绘专业的研究生（包括助教进修生）的教材，也可供本科高年级学生和测绘科技工作者参考。

编著者

1991年7月于武汉

目　　录

第一章　随机过程基础…………………………………………………………………（ 1 ）
§1-1　多维正态分布…………………………………………………………………（ 1 ）
§1-2　随机过程的概率分布和数字特征……………………………………………（ 6 ）
§1-3　平稳随机过程…………………………………………………………………（ 11 ）
§1-4　协方差函数与相关函数………………………………………………………（ 15 ）
§1-5　几个特殊的随机过程…………………………………………………………（ 19 ）
第二章　估计方法和广义测量平差原理…………………………………………………（ 23 ）
§2-1　概述……………………………………………………………………………（ 23 ）
§2-2　极大似然估计…………………………………………………………………（ 25 ）
§2-3　最小二乘估计…………………………………………………………………（ 28 ）
§2-4　极大验后估计…………………………………………………………………（ 31 ）
§2-5　最小方差估计…………………………………………………………………（ 34 ）
§2-6　线性最小方差估计……………………………………………………………（ 37 ）
§2-7　贝叶斯（Bayes）估计…………………………………………………………（ 39 ）
§2-8　广义测量平差原理……………………………………………………………（ 42 ）
第三章　秩亏自由网平差…………………………………………………………………（ 46 ）
§3-1　平差问题的基准与秩亏自由网平差的类型…………………………………（ 46 ）
§3-2　秩亏自由网平差………………………………………………………………（ 51 ）
§3-3　秩亏水准网的动态平差………………………………………………………（ 58 ）
§3-4　秩亏自由网的分区平差………………………………………………………（ 63 ）
§3-5　附加系统参数的秩亏自由网平差……………………………………………（ 72 ）
§3-6　秩亏自由网平差的特点………………………………………………………（ 78 ）
第四章　平差随机模型的验后估计………………………………………………………（ 85 ）
§4-1　概述……………………………………………………………………………（ 85 ）
§4-2　赫尔默特估计法………………………………………………………………（ 86 ）
§4-3　赫尔默特估计的一些简化公式………………………………………………（102）
§4-4　方差-协方差估计法 …………………………………………………………（110）
§4-5　二次无偏估计法………………………………………………………………（119）
§4-6　方差分量估计中的精度评定…………………………………………………（129）
第五章　滤波与配置………………………………………………………………………（138）
§5-1　概述……………………………………………………………………………（138）
§5-2　极大验后滤波与推估…………………………………………………………（139）

§5-3 最小二乘滤波与推估……………………………………………………（146）
§5-4 最小二乘配置……………………………………………………………（152）
§5-5 滤波与配置的验后方差…………………………………………………（163）
§5-6 静态逐次滤波……………………………………………………………（167）
第六章 若干平差问题的补充……………………………………………（175）
§6-1 最小二乘平差中部分主子块秩亏时的求解方法………………………（175）
§6-2 最小二乘平差中各种主要协因数阵的秩………………………………（181）
§6-3 顾及起算数据误差时的精度评定………………………………………（184）
§6-4 随机模型具有奇异协因数阵的平差……………………………………（189）
§6-5 向量空间投影与最小二乘原理…………………………………………（194）
§6-6 向量空间理论中的平差问题……………………………………………（198）
§6-7 广义 G-M 模型下的平差问题 …………………………………………（202）
§6-8 广义 G-M 模型下的精度和统计性质 …………………………………（207）
§6-9 等价观测值………………………………………………………………（213）
第七章 动态线性系统的卡尔曼滤波……………………………………（219）
§7-1 连续线性系统的数学模型………………………………………………（219）
§7-2 离散线性系统的数学模型………………………………………………（228）
§7-3 离散线性系统的卡尔曼滤波……………………………………………（234）
§7-4 用正交投影法推导卡尔曼滤波…………………………………………（243）
§7-5 卡尔曼滤波的应用举例…………………………………………………（248）
§7-6 离散型卡尔曼滤波的推广………………………………………………（255）
§7-7 离散线性系统的预测……………………………………………………（260）
§7-8 离散线性系统的平滑……………………………………………………（264）
§7-9 连续线性系统的卡尔曼滤波……………………………………………（270）
第八章 卡尔曼滤波的渐近性质和误差分析……………………………（274）
§8-1 线性确定系统的能观性和能控性………………………………………（274）
§8-2 卡尔曼滤波的稳定性……………………………………………………（278）
§8-3 模型误差分析……………………………………………………………（283）
§8-4 滤波的发散现象和克服发散的方法……………………………………（287）
附录…………………………………………………………………………（292）
§1 矩阵的秩……………………………………………………………………（292）
§2 方阵的迹……………………………………………………………………（292）
§3 矩阵的正定性和许瓦茨不等式……………………………………………（295）
§4 Dirac-δ 函数 ………………………………………………………（298）
§5 拉普拉斯变换………………………………………………………………（299）
§6 向量空间的概念……………………………………………………………（301）
主要参考文献………………………………………………………………（305）

第一章　随机过程基础

本章在具有一定的概率论和数理统计知识的基础上，介绍随机过程的基础知识，目的为了以后应用。

§1-1　多维正态分布

一、多维正态分布的定义和性质

设有m个互相独立的标准正态随机变量构成的随机向量 $Z=[Z_1\ Z_2\ \cdots\ Z_m]^T$，则称它们的有限个线性函数

$$\underset{n\times 1}{X}=\begin{pmatrix}X_1\\X_2\\\vdots\\X_n\end{pmatrix}=\underset{n\times m}{A}\begin{pmatrix}Z_1\\Z_2\\\vdots\\Z_m\end{pmatrix}+\underset{n\times 1}{\mu} \tag{1-1-1}$$

为n维正态随机向量。此时，X的数学期望和方差阵为

$$\left.\begin{aligned}E(X)&=\mu\\D_X&=AA^T\end{aligned}\right\} \tag{1-1-2}$$

X的分布函数和概率密度都简称为n维（或n元）正态分布，简记为 $X\sim N_n(\mu,AA^T)$，或写为 $X\sim N(\mu,\ D_X)$。

由互相独立的标准正态随机变量组成的随机向量Z，可写为$Z\sim N(0,\ E_n)$。E_n为n阶单位阵。

多维正态分布具有以下性质：

（1）正态随机向量的线性函数还是正态的，例如，设 $X\sim N_n(\mu,\ AA^T)$，$Y=BX+b$，则

$$Y\sim N(B\mu+b,\ BAA^TB^T).$$

（2）设 $X\sim N_n(\mu,\ AA^T)$，记

$$X=\begin{bmatrix}X_1\\X_2\end{bmatrix},\ \mu=\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix},\ AA^T=\begin{bmatrix}D_{11}&D_{12}\\D_{21}&D_{22}\end{bmatrix},$$

则$\underset{r\times 1}{X_1}\sim N_r(\mu_1,\ D_{11})$，$X_2\sim\underset{n-r}{N}(\mu_2,\ D_{22})$。

二、多维正态分布

设有n维正态随机向量$N_n(\mu_X,\ D_X)$，其中方差阵D_X为可逆阵，即$\det(D_X)\neq 0$，则它的概率密度为

$$f(x)=(2\pi)^{-\frac{n}{2}}\,|D_X|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x-\mu_X)^T D_X^{-1}(x-\mu_X)\right\}. \tag{1-1-3}$$

式中$|D_X|$表示D_X的行列式。

对于二维正态随机向量$[X\ \ Y]^T$，若它有可逆方差阵和数学期望为

$$\begin{bmatrix}\sigma_X^2 & \sigma_{XY}\\ \sigma_{XY} & \sigma_Y^2\end{bmatrix}\text{和}\begin{bmatrix}\mu_X\\ \mu_Y\end{bmatrix},$$

则由（1-1-3）式可得其概率密度为

$$f(x,\ y)=\frac{1}{2\pi\sqrt{\sigma_X^2\sigma_Y^2-\sigma_{XY}^2}}\exp\left\{-\frac{(x-\mu_X)^2\sigma_Y^2-2(x-\mu_X)(y-\mu_Y)\sigma_{XY}+(y-\mu_Y)^2\sigma_X^2}{2(\sigma_X^2\sigma_Y^2-\sigma_{XY}^2)}\right\};$$

因相关系数 $\rho_{XY}=\dfrac{\sigma_{XY}}{\sigma_X\sigma_Y}$，所以上式可写为

$$f(x,\ y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}}\exp\left\{-\frac{1}{2(1-\rho_{XY}^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]\right\}, \tag{1-1-4}$$

这就是二维正态随机向量概率密度。

当 $\rho_{XY}=0$ 或 $\sigma_{XY}=0$ 时，即当X和Y是互不相关的两个正态随机变量时，则有

$$\begin{aligned}f(x,\ y)&=\frac{1}{2\pi\sigma_X\sigma_Y}\exp\left\{-\frac{(x-\mu_X)^2}{2\sigma_X^2}-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right\}\\&=\frac{1}{\sqrt{2\pi}\,\sigma_X}\exp\left\{-\frac{(x-\mu_Y)^2}{2\sigma_X^2}\right\}\cdot\frac{1}{\sqrt{2\pi}\,\sigma_Y}\exp\left\{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right\}\\&=f_x(x)f_y(y).\end{aligned} \tag{1-1-5}$$

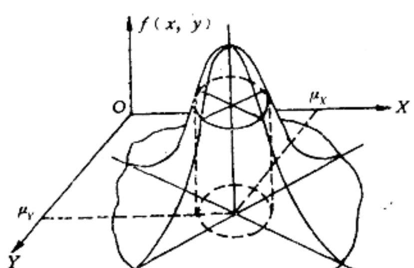

图 1-1

这就是说，当 $\rho_{XY}=0$ 时，X和Y是互相独立的。所以，对于正态分布来说，随机变量的“互不相关”与“互相独立”是等价的。

根据（1-1-4）式绘制二维正态曲面（密度曲面）如图 1-1。曲面在点（μ_x，μ_y）处取得最大值。如果用平行于 XOY 面的平面 $Z=Z_0$（常数）截此曲面，即得到一族椭圆，椭圆上所有点的概率密度值均相等，因此，称这些椭圆为等密度椭圆。

三、正态随机向量的条件概率密度

设有 $n+t$ 维正态随机向量X，且设

$$X=\begin{bmatrix}X_1\\ X_2\end{bmatrix},\quad \mu_X=\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix},\quad D_X=\begin{bmatrix}D_{11} & D_{12}\\ D_{21} & D_{22}\end{bmatrix},$$

X_1和X_2分别是由X的前 n 个分量和后 t 个分量构成的正态随机向量，即 $X_1\sim N_n(\mu_1,$

D_{11})，$X_2 \sim N_t(\mu_2, D_{22})$。$X$的概率密度是

$$f(x)=(2\pi)^{-\frac{s+t}{2}}\,|D_X|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}^T D_X^{-1}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}\right\}. \qquad (1\text{-}1\text{-}6)$$

按分块矩阵求逆公式，有

$$D_X^{-1}=\begin{bmatrix} D_{11}^{-1}+D_{11}^{-1}D_{12}\widetilde{D}_{22}^{-1}D_{21}D_{11}^{-1} & -D_{11}^{-1}D_{12}\widetilde{D}_{22}^{-1} \\ -\widetilde{D}_{22}^{-1}D_{21}D_{11}^{-1} & \widetilde{D}_{22}^{-1} \end{bmatrix}, \qquad (1\text{-}1\text{-}7a)$$

或为

$$D_X^{-1}=\begin{bmatrix} \widetilde{D}_{11}^{-1} & -\widetilde{D}_{11}^{-1}D_{12}D_{22}^{-1} \\ -D_{22}^{-1}D_{21}\widetilde{D}_{11}^{-1} & D_{22}^{-1}+D_{22}^{-1}D_{21}\widetilde{D}_{11}^{-1}D_{12}D_{22}^{-1} \end{bmatrix}, \qquad (1\text{-}1\text{-}7b)$$

其中

$$\left.\begin{aligned} \widetilde{D}_{11}&=D_{11}-D_{12}D_{22}^{-1}D_{21} \\ \widetilde{D}_{22}&=D_{22}-D_{21}D_{11}^{-1}D_{12} \end{aligned}\right\}, \qquad (1\text{-}1\text{-}7c)$$

可将（1-1-7a）和（1-1-7b）两式分别写为

$$D_X^{-1}=\begin{bmatrix} -D_{11}^{-1}D_{12} \\ E_t \end{bmatrix}\widetilde{D}_{22}^{-1}\begin{bmatrix} -D_{21}D_{11}^{-1} & E_t \end{bmatrix}+\begin{bmatrix} D_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}, \qquad (1\text{-}1\text{-}7d)$$

$$D_X^{-1}=\begin{bmatrix} E_s \\ -D_{22}^{-1}D_{21} \end{bmatrix}\widetilde{D}_{11}^{-1}\begin{bmatrix} E_s & -D_{12}D_{22}^{-1} \end{bmatrix}+\begin{bmatrix} 0 & 0 \\ 0 & D_{22}^{-1} \end{bmatrix}, \qquad (1\text{-}1\text{-}7e)$$

因D_X还可分解为

$$D_X=\begin{bmatrix} D_{11} & 0 \\ D_{21} & \widetilde{D}_{22} \end{bmatrix}\begin{bmatrix} E & D_{11}^{-1}D_{12} \\ 0 & E \end{bmatrix}=\begin{bmatrix} \widetilde{D}_{11} & D_{12} \\ 0 & D_{22} \end{bmatrix}\begin{bmatrix} E & 0 \\ D_{22}^{-1}D_2 & E \end{bmatrix}, \qquad (1\text{-}1\text{-}8a)$$

所以，D_X的行列式之值为

$$|D_X|=|D_{11}|\,|\widetilde{D}_{22}|=|D_{22}|\,|\widetilde{D}_{11}|. \qquad (1\text{-}1\text{-}8b)$$

利用(1-1-7d)、(1-1-7c)和(1-1-8b)式，可将概率密度式(1-1-6)改写为

$$\begin{aligned} f(x)&=f(x_1,\ x_2) \\ &=(2\pi)^{-\frac{s}{2}}|D_{11}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x_1-\mu_1)^T D_{11}^{-1}(x_1-\mu_1)\right\}. \\ &\quad\cdot(2\pi)^{-\frac{t}{2}}|\widetilde{D}_{22}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x_2-\widetilde{\mu}_2)^T\widetilde{D}_{22}^{-1}(x_2-\widetilde{\mu}_2)\right\}, \end{aligned} \qquad (1\text{-}1\text{-}9a)$$

或

$$\begin{aligned} f(x)&=f(x_1,\ x_2) \\ &=(2\pi)^{-\frac{t}{2}}|D_{22}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x_2-\mu_2)^T D_{22}^{-1}(x_2-\mu_2)\right\} \\ &\quad\cdot(2\pi)^{-\frac{s}{2}}|\widetilde{D}_{11}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x_1-\widetilde{\mu}_1)^T\widetilde{D}_{11}^{-1}(x_1-\widetilde{\mu}_1)\right\}, \end{aligned} \qquad (1\text{-}1\text{-}9b)$$

其中

$$\left.\begin{aligned}\tilde{\mu}_1 &= \mu_1 + D_{12}D_{22}^{-1}(x_2 - \mu_2)\\ \tilde{\mu}_2 &= \mu_2 + D_{21}D_{11}^{-1}(x_1 - \mu_1)\end{aligned}\right\},\tag{1-1-9c}$$

根据边际概率密度和多维正态分布的性质可知

$$f_1(x_1) = (2\pi)^{-\frac{s}{2}}|D_{11}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x_1 - \mu_1)^T D_{11}^{-1}(x_1 - \mu_1)\right\},\tag{1-1-10a}$$

$$f_2(x_2) = (2\pi)^{-\frac{t}{2}}|D_{22}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x_2 - \mu_2)^T D_{22}^{-1}(x_2 - \mu_2)\right\}.\tag{1-1-10b}$$

又由条件概率密度公式知

$$f(x_2/x_1) = \frac{f(x_1,\ x_2)}{f_1(x_1)},\tag{1-1-11a}$$

$$f(x_1/x_2) = \frac{f(x_1,\ x_2)}{f_2(x_2)}.\tag{1-1-11b}$$

将（1-1-9a）和（1-1-10a）两式代入（1-1-11a）式，得

$$f(x_2/x_1) = (2\pi)^{-\frac{t}{2}}|\tilde{D}_{22}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x_2 - \tilde{\mu}_2)^T\tilde{D}_{22}^{-1}(x_2 - \tilde{\mu}_2)\right\},\tag{1-1-12a}$$

而将（1-1-9b）和（1-1-10b）两式代入（1-1-11b）式，即得

$$f(x_1/x_2) = (2\pi)^{-\frac{s}{2}}|\tilde{D}_{11}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x_1 - \tilde{\mu}_1)^T\tilde{D}_{11}^{-1}(x_1 - \tilde{\mu}_1)\right\}.\tag{1-1-12b}$$

显然，上两式仍然是正态概率密度，根据条件期望和条件方差的定义和正态概率密度的性质可得

$$\left.\begin{aligned}E(X_1/x_2) &= \tilde{\mu}_1 = \mu_1 + D_{12}D_{22}^{-1}(x_2 - \mu_2)\\ E(X_2/x_1) &= \tilde{\mu}_2 = \mu_2 + D_{21}D_{11}^{-1}(x_1 - \mu_1)\end{aligned}\right\},\tag{1-1-12c}$$

$$\left.\begin{aligned}D(X_1/x_2) &= \tilde{D}_{11} = D_{11} - D_{12}D_{22}^{-1}D_{21}\\ D(X_2/x_1) &= \tilde{D}_{22} = D_{22} - D_{21}D_{11}^{-1}D_{12}\end{aligned}\right\}.\tag{1-1-12d}$$

因此，(1-1-12a)和(1-1-12b)又可写为

$$\left.\begin{aligned}f(x_2/x_1) &= (2\pi)^{-\frac{t}{2}}|D(X_2/x_1)|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}[x_2 - E(X_2/x_1)]^T\right.\\ &\qquad\left.\cdot D^{-1}(X_2/x_1)[x_2 - E(X_2/x_1)]\right\}\\ f(x_1/x_2) &= (2\pi)^{-\frac{s}{2}}|D(X_1/x_2)|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}[x_1 - E(X_1/x_2)]^T\right.\\ &\qquad\left.\cdot D^{-1}(X_1/x_2)[x_1 - E(X_1/x_2)]\right\}\end{aligned}\right\}.\tag{1-1-13}$$

正态分布的条件期望具有以下性质：

（1）由（1-1-12c）式可知，$E(X_1/x_2)$是x_2的线性组合，所以，它是正态随机向量；当然，$E(X_2/x_1)$也是正态随机向量。

（2）设X和Y为正态随机向量，且设

$$\left.\begin{aligned}\widetilde{X}&=X-E(X/y)\\ Z&=AY\end{aligned}\right\},\tag{1-1-14}$$

则$\widetilde{X}$是与Z互相独立的随机向量。这是因为

$$\widetilde{X}=X-\{\mu_X+D_{XY}D_Y^{-1}(Y-\mu_Y)\}=X-D_{XY}D_Y^{-1}Y+D_{XY}D_Y^{-1}\mu_Y,$$

由协方差传播律可得

$$D(\widetilde{X},Z)=[E\quad -D_{XY}D_Y^{-1}]\begin{bmatrix}D_X & D_{XY}\\ D_{YX} & D_Y\end{bmatrix}\begin{bmatrix}0\\ A^T\end{bmatrix}$$

$$=D_{XY}A^T-D_{XY}D_Y^{-1}D_YA^T=0.$$

（3）设$X\sim N(\mu_X,D_X)$，$Y_1\sim N(\mu_1,D_1)$，$Y_2\sim N(\mu_2,D_2)$，且$\mathrm{cov}(Y_1,Y_2)=0$，而$D(X,Y_1)=D_{XY_1}\neq 0$，$D(X,Y_2)=D_{XY_2}\neq 0$，则有

$$E(X/y)=E(X/y_1,y_2)=E(X/y_1)+E(X/y_2)-\mu_X.\tag{1-1-15}$$

证：因为

$$E(X/y)=E(X/y_1,y_2)=\mu_X+D_{XY}D_Y^{-1}(y-\mu_Y)$$

$$=\mu_X+[D_{XY_1}\quad D_{XY_2}]\begin{bmatrix}D_1^{-1} & 0\\ 0 & D_2^{-1}\end{bmatrix}\begin{bmatrix}y_1-\mu_1\\ y_2-\mu_2\end{bmatrix},$$

所以

$$E(X/y_1,y_2)=\{\mu_X+D_{XY_1}D_1^{-1}(y_1-\mu_1)\}+\{D_{XY_2}D_2^{-1}(y_2-\mu_2)+\mu_X-\mu_X\}$$

$$=E(X/y_1)+E(X/y_2)-\mu_X.$$

（4）设$X\sim N(\mu_X,D_X)$，$Y\sim N(\mu_Y,D_Y)$，且

$$\mu_Y=\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix},\quad D_Y=\begin{bmatrix}D_{11} & D_{12}\\ D_{21} & D_{22}\end{bmatrix},\quad D_{YX}=\begin{bmatrix}D_{Y_1X}\\ D_{Y_2X}\end{bmatrix}=D_{XY}^T,$$

令$\widetilde{Y}_2=Y_2-E(Y_2/y_1)$，则有

$$E(X/y_1,y_2)=E(X/y_1,\widetilde{y}_2)=E(X/y_1)+E(X/\widetilde{y}_2)-\mu_X.\tag{1-1-16}$$

证：因为

$$\widetilde{Y}_2=Y_2-\mu_2-D_{21}D_{11}^{-1}(Y_1-\mu_1)$$

$$=[-D_{21}D_{11}^{-1}\quad E]\begin{bmatrix}Y_1\\ Y_2\end{bmatrix}-\mu_2+D_{21}D_{11}^{-1}\mu_1$$

所以

$$\left.\begin{aligned}&E(\widetilde{Y}_2)=0,\\ &D(\widetilde{Y}_2)=[-D_{21}D_{11}^{-1}\quad E]\begin{bmatrix}D_{11} & D_{12}\\ D_{21} & D_{22}\end{bmatrix}\begin{bmatrix}-D_{11}^{-1}D_{12}\\ E\end{bmatrix}\\ &\qquad=D_{22}-D_{21}D_{11}^{-1}D_{12}=\widetilde{D}_{22}\\ &D(\widetilde{Y}_2,Y_1)=0,\\ &D(\widetilde{Y}_2,X)=[-D_{21}D_{11}^{-1}\quad E]\begin{bmatrix}D_{Y_1X}\\ D_{Y_2X}\end{bmatrix}\\ &\qquad=D_{Y_2X}-D_{21}D_{11}^{-1}D_{Y_1X}.\end{aligned}\right\}\tag{1-1-17}$$

利用分块求逆公式和（1-1-17）式得

$$
\begin{aligned}
E(X/y_1,\ y_2) &= \mu_X + D_{XY}D_Y^{-1}(Y-\mu_Y) \\
&= \mu_X + [D_{XY_1}\quad D_{XY_2}]\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22}\end{bmatrix}^{-1}\begin{bmatrix} Y_1-\mu_1 \\ Y_2-\mu_2\end{bmatrix} \\
&= \mu_X + [D_{XY_1}\quad D_{XY_2}]\left\{\begin{bmatrix} -D_{11}^{-1}D_{12} \\ E\end{bmatrix}\widetilde{D}_{22}^{-1}[-D_{21}D_{11}^{-1}\quad E] \right. \\
&\quad \left. + \begin{bmatrix} D_{11}^{-1} & 0 \\ 0 & 0\end{bmatrix}\right\}\begin{bmatrix} Y_1-\mu_1 \\ Y_2-\mu_2\end{bmatrix} \\
&= \mu_X + \{(-D_{XY_1}D_{11}^{-1}D_{12} + D_{XY_2})D^{-1}(\widetilde{Y}_2)[-D_{21}D_{11}^{-1}\quad E] \\
&\quad + [D_{XY_1}D_{11}^{-1}\quad 0]\}\begin{bmatrix} Y_1-\mu_1 \\ Y_2-\mu_2\end{bmatrix} \\
&= \mu_X + D(X,\widetilde{Y}_2)D^{-1}(\widetilde{Y}_2)\{-D_{21}D_{11}^{-1}(Y_1-\mu_1)+Y_2-\mu_2\} \\
&\quad + D_{XY_1}D_{11}^{-1}(Y_1-\mu_1)+\mu_X-\mu_X \\
&= E(X/y_1)+E(X/\widetilde{y}_2)-\mu_X.
\end{aligned}
$$

§1-2　随机过程的概率分布和数字特征

一、随机过程的基本概念

随机变量是在一个实验结果中可能取得某个数值，但不能预先知道取何值的变量。它是不随时间或其它因素而变化的。在自然科学和工程技术的许多实际问题中，往往还存在一类时间或其它因素（如位置等）而不断变化的随机现象，由此构成了一族无穷多个互相关联的随机变量。通常用 t 表示时间（或其它某种因素），而称这种依赖于 t 的一族无穷多个互相关联的随机变量之集合为随机过程，或称为随机函数。换句话说，随机过程是这样的函数，它是时间 t 的函数，且能在实验的结果中取某种确定的但预先未知的形式。一般用 $X(t)$ 表示随机过程，有时也记为 $\{X(t),\ t\in T\}$。对于一个固定的 t_i，$X(t_i)$ 就是一个随机变量。显然，随机过程是一个由随机变量扩展出来的概念。

随机过程在实验结果中所取的具体形式称为这个随机过程的实现，或称为样本函数。通常用 $x(t)$ 表示 $X(t)$ 的“实现”。一次实验或一次观测得到的结果，就是 $X(t)$ 的一个“实现”。对随机过程的一组试验，就得到它的一组“实现”。

这里先举一个随机过程的例子。

有某金属丝厂制造一种长度为 l，断面直径为 0.2mm 的细金属丝，虽然制造的条件相同，但由于一些随机因素的影响，使得制造出来的每根金属丝上各点处的断面直径不可能正好是 0.2mm，而是在 0.2mm 附近摆动，并且随各点至起点（一个端点）的距离的不同而变化。或者说是距离的一种预先未知其形式的函数，因此，金属丝的断面直径是一个随机过程。我们也用 $X(t)$ 表示这个随机过程。其中 t 是金属丝上各点至起点的距离。为了便于说明，姑且用一条曲线来表示 $X(t)$，如图 1-2（严格说来，不能在图上用曲线表示）。由该图可以看到，若以确定的 t_i 值代入 $X(t)$，所得的 $X(t_i)(i=1,\ 2,\ \cdots\cdots,\ n)$ 都

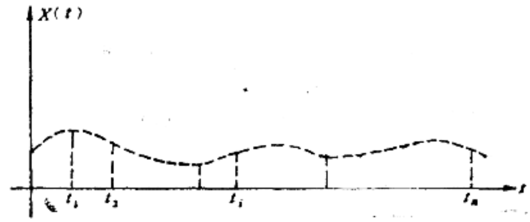

图 1-2

是一般的随机变量。若用 $X(t_1)$、$X(t_2)$、…，$X(t_n)$代替$X(t)$，则个数 n 愈多，这种代替愈精确。制造的各根金属丝的断面直径可以直接精确测定，测出一根金属丝上几个点处的断面直径就得到这个随机过程的一个实现，因此，制造得到一根金属丝，就可以看成得到一个实现，得到另一根金属丝，就是得到另一个实现。图 1-3 给出了编号为 1 、 2 、 3 和 4 号的四根金属丝的实现 $x_1(t)$、$x_2(t)$、$x_3(t)$和$x_4(t)$。

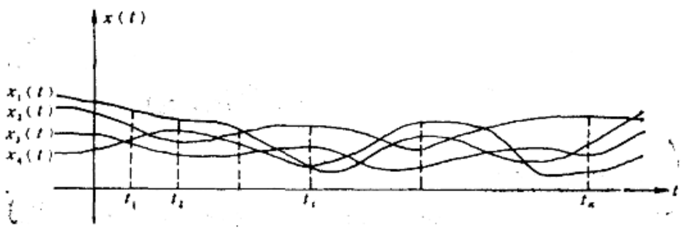

图 1-3

还可以举一些例子。

测定人造地球卫星到地面站的距离，对于固定的时刻 t，这个距离是一个随机变量。随着时刻 t 的变化，测定的距离就构成一个随机过程。卫星通过观测站的上空时所测定的距离就是它的一个实现。

航测像片上网格点的位置也是一个随机过程。因为像片上各网格点的位置可以测定。所以，每一张像片就是随机过程的一个实现。

电话总机接到的呼唤次数具有随机性，若以某个时刻为零时刻，以 $X(t)$表示时刻 t 以前已经接到的呼唤次数，对于固定的时刻 t，它是一个随机变量，因此，随着 t 的变化，就得到一个随机过程。

随机过程可以分成两类，如果考虑的因素 t（时间，距离等）是离散的，则称这种随机过程为随机序列，如航测像片的网格点；如果考虑的因素 t 是连续的，则仍称它随机过程，如金属丝的断面直径、卫星到观测站的距离和电话问题。以后主要讨论 t 是连续的随机过程。并假定 t 是表示时间，当然 t 也可以是其它因素。

二、随机过程的概率分布

随机过程在确定的时刻上是一个随机变量，因此，可以利用随机变量的统计描述方法

来描述随机过程的统计特性。

设 $X(t)$是一个随机过程，对于一个固定的 t_1，$X(t_1)$就是一个随机变量。$X(t_1)<x_1$ 的概率为 $P(X(t_1)<x_1)$，它是 x_1和 t 的函数，记作

$$F_1(x_1, t_1)=P(X(t_1)<x_1), \tag{1-2-1a}$$

称它为随机过程 $X(t)$的一维分布函数。

如果 $F_1(x_1, t)$对 x_1的偏导数存在，则有

$$f_1(x_1, t_1)=\frac{\partial F_1(x_1, t_1)}{\partial x_1}, \tag{1-2-1b}$$

称 $f_1(x_1, t_1)$为随机过程 $X(t)$的一维概率密度。

随机过程的一维分布函数和一维概率密度给出了随机过程最简单的概率分布特性，它们描述了随机过程在各个孤立时刻的统计特性，但不能反映出随机过程在不同时刻的联系。为此，还应当考虑各个不同时刻的情况，最简单的情况是 $t=t_1$和 $t=t_2$时的二维分布。

随机过程 $X(t)$在 t_1、t_2两个时刻的随机变量 $X(t_1)$和 $X(t_2)$的联合分布函数，记为

$$F_2(x_1, x_2; t_1, t_2)=P\{X(t_1)<x_1; X(t_2)<x_2\}, \tag{1-2-2a}$$

称为随机过程 $X(t)$的二维分布函数。如果它对 x_1、x_2的二阶偏导数存在，有

$$f_2(x_1, x_2; t_1, t_2)=\frac{\partial^2 F_2(x_1, x_2; t_1, t_2)}{\partial x_1 \partial x_2}, \tag{1-2-2b}$$

称它为随机过程的二维概率密度。

用同样的方法可以建立 $X(t)$的 n 维分布函数和 n 维概率密度为

$$\left.\begin{aligned}&F_n(x_1, x_2, \cdots\cdots, x_n; t_1, t_2, \cdots\cdots, t_n)\\&=P\{X((t_1)<x_1, X(t_2)<x_2, \cdots\cdots, X(t_n)<x_n)\}\\&f_n(x_1, x_2, \cdots\cdots, x_n; t_1, t_2, \cdots\cdots, t_n)\\&=\frac{\partial^n F_n(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n)}{\partial x_1 \partial x_2 \cdots\cdots \partial x_n}\end{aligned}\right\} \tag{1-2-3}$$

三、随机过程的数字特征

可以看到，从理论上说，知道了随机过程 $X(t)$的 n 维概率分布就能知道其全部统计特性。但在实际问题中，要给出随机过程的分布函数和概率密度一般是不可能的，也是不必要的。因此，也可以利用随机过程的一些数字特征来描述它的基本统计性质。这些数字特征都是时间 t 的函数，主要有：

1. 数学期望。对于每个确定的 t，由随机过程所确定的随机变量的数学期望

$$\mu_X(t)=E(X(t))=\int_{-\infty}^{\infty} x f(x, t) dx. \tag{1-2-4}$$

称为 $X(t)$的数学期望，也称为均值函数。它是一个普通的数值函数，是随机过程 $X(t)$的所有样本函数在时刻 t 的函数值的平均。

2. 方差。对于每个确定时刻 t，由随机过程所确定的随机变量的方差

$$D_X(t)=\mathrm{var}(X(t))=E\{(X(t))-\mu_X(t))^2\}$$
$$=\int_{-\infty}^{\infty}(x(t)-\mu_X(t))^2 f(x,\ t)dx, \qquad (1\text{-}2\text{-}5)$$

称为 $X(t)$的方差。它是依赖于 t 的非负函数，描述了随机过程的各个样本对于数学期望的离散程度。$D_X(t)$的平方根称为随机过程的中误差(或标准差)，记为

$$\sigma_X(t)=\sqrt{D_X(t)}. \qquad (1\text{-}2\text{-}6)$$

3. 自协方差函数和自相关函数，对于任意时刻 t_1、t_2随机变量 $X(t_1)$和 $X(t_2)$有协方差

$$D_X(t_1,\ t_2)=\mathrm{cov}(X(t_1),\ X(t_2))$$
$$=E\{(X(t_1)-\mu_X(t_1))(X(t_2)-\mu_X(t_2))\}, \qquad (1\text{-}2\text{-}7)$$

则称 $D_X(t_1,\ t_2)$为随机过程的自协方差函数，简称协方差函数。当 $t=t_1=t_2$时，协方差函数就是随机过程的方差 $D_X(t)$。因此，协方差函数是研究随机过程的主要概念。

在时刻 t_1、t_2，$X(t_1)$和 $X(t_2)$也有原点相关矩

$$R_X(t_1,\ t_2)=E\{X(t_1)X(t_2)\}, \qquad (1\text{-}2\text{-}8)$$

称 $R_X(t_1,\ t_2)$为随机过程 $X(t)$的自相关函数，简称为相关函数。

协方差函数和相关函数有关系：

$$D_X(t_1,\ t_2)=R_X(t_1,\ t_2)-\mu_X(t_1)\mu_X(t_2), \qquad (1\text{-}2\text{-}9)$$

[例 1-2-1] 已知随机过程 $X(t)=\xi\sin t-\eta\cos t$,其中 ξ、η 是相互独立，且服从 $N(0,\ \sigma^2)$ 的正态随机变量，求 $X(t)$的数学期望、方差、自协方差函数和相关函数。

解：
$$\mu_X(t)=E(\xi)\sin t+E(\eta)\cos\xi=0,$$
$$D_X(t)=\sin^2 tD_\xi+\cos^2 tD_\eta=\sigma^2$$
$$D_X(t_1,\ t_2)=R_X(t_1,\ t_2)=D(X(t_1),\ X(t_2))$$
$$=E\{(\xi\sin t_1+\eta\cos t_1)(\xi\sin t_2+\eta\cos t_2)\}$$
$$=\sin t_1\sin t_2\sigma^2+\cos t_1\cos t_2\sigma^2$$
$$=\sigma^2\cos(t_1-t_2).$$

4. 互协方差函数和互相关函数。当考虑两个或两个以上的随机过程时，还需要用互协方差函数或互相关函数来描述它们之间的相关关系。随机过程 $X(t)$和 $Y(t)$之间的互协方差函数和互相关函数的定义是

$$D_{XY}(t_1,\ t_2)=E\{[X(t_1)-\mu_X(t_1)][Y(t_2)-\mu_Y(t_2)]\}$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[x(t_1)-\mu_X(t_1)][y(t_2)-\mu_Y(t_2)]$$
$$\cdot f_z(x,\ t_1;\ y,\ t_2)d_xd_y \qquad (1\text{-}2\text{-}10a)$$
$$R_{XY}(t_1,\ t_2)=E\{X(t_1)Y(t_2)\}$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x(t_1)y(t_2)f_z(x,\ t_1;\ y,\ t_2)d_xd_y \qquad (1\text{-}2\text{-}10b)$$

式中 $f_z(x,\ y;\ t_1,\ t_2)$是 $X(t)$和 $Y(t)$二维联合概率密度。

有时还定义标准化自相关函数和标准化互相关函数为

$$\rho_X(t_1,\ t_2)=\frac{D_X(t_1,\ t_2)}{\sigma(t_1)\sigma(t_2)} \tag{1-2-11a}$$

$$\rho_{XY}(t_1,\ t_2)=\frac{D_{XY}(t_1,\ t_2)}{\sigma_X(t_1)\sigma_Y(t_2)} \tag{1-2-11b}$$

当 t_1 和 t_2 固定时，$\rho_X(t_1,\ t_2)$ 就是 $X(t_1)$ 和 $X(t_2)$ 的相关系数，$\rho_{XY}(t_1,\ t_2)$ 就是 $X(t_1)$ 和 $Y(t_2)$ 的相关系数。

四、随机向量过程

如果考虑组成随机向量的多个随机变量，它们都随 t 的不同而变化，这就同时得到了多个随机过程，称之为随机向量过程，其中离散型的也称为随机向量序列。

例如，观测人造卫星的状态(位置和速度)，在某个固定的时刻 t，卫星的位置和速度是一个随机向量。而当 t 不断变化时，就构成了一个随机向量过程。

设 $X(t)$ 是一个由随机过程 $X_1(t)$，$X_2(t)$，……，$X_n(t)$ 组成的随机向量过程，即可写为

$$X(t)=[X_1(t),\ X_2(t),\ \cdots,\ X_n(t)]^T.$$

$X(t)$ 也是 t 的函数。

显然，也可以用描述随机过程的概率分布及其数字特征来描述随机向量过程。此时，只要将上面的(1-2-1)～(1-2-5)、(1-2-7)～(1-2-10)诸式作适当的推广，将随机过程变换为随机向量过程就行。

对于随机过程或随机向量过程 $X(t)$，如果对任意时间点 t_1、$t_2\cdots t_m$，有

$$f(x_1,\ x_2,\cdots\cdots,x_m;\ t_1,\ t_2,\ \cdots\cdots,\ t_m)=\prod_{i=1}^{m} f_i(x_i,\ t_i), \tag{1-2-12}$$

则称 $X(t)$ 是独立随机过程或独立随机向量过程。

设有随机向量过程 $X(t)$ 和 $Y(t)$；如果对任意时间点 t_1，t_2，……，t_m，有

$$\begin{aligned}&f(x_1,x_2,\cdots\cdots,\ x_m,\ t_1,\ t_2,\ \cdots\cdots,\ t_m;\ y_1,\ y_2,\ \cdots\cdots,y_m,t_1,\\&\quad t_2,\cdots\cdots,t_m)=f_1(x_1,\ x_2,\cdots\cdots,x_m;\ t_1,t_2,\cdots\cdots,t_m)\cdot f_2(y_1,\\&\quad y_2,\cdots\cdots,y_m;\ t_1,\ t_2,\ \cdots\cdots,\ t_m),\end{aligned} \tag{1-2-13}$$

则称 $X(t)$ 和 $Y(t)$ 是两个互相独立的随机向量过程。

对于随机向量过程 $X(t)$，若所有的 t_1，$t_2(t_1\neq t_2)$ 均有

$$R_X(t_1,\ t_2)=\mu_X(t_1)\mu_X^T(t_2), \tag{1-2-14a}$$

或

$$D_X(t_1,\ t_2)=0, \tag{1-2-14b}$$

则称 $X(t)$ 是不相关随机向　过程。

对于随机向量过程 $X(t)$ 和 $Y(t)$，若对所有的 t_1，t_2，均有

$$R_{XY}(t_1,\ t_2)=\mu_x(t_1)\mu_Y^T(t_2), \tag{1-2-15a}$$

或

$$D_{XY}(t_1,\ t_2)=0, \tag{1-2-15b}$$

则称$X(t)$和$Y(t)$为彼此不相关的随机向量过程。

容易理解，一个独立的随机向量过程是一个不相关的过程，但不相关的过程不一定是一个独立过程；同样，两个互相独立的随机向量过程也是彼此不相关的，但反之也不一定成立。

§1-3 平稳随机过程

一、平稳随机过程的基本概念

在实际问题中，经常用到的一类重要的随机过程是平稳随机过程，它的特点是其统计性质与所选取的时间t的起点无关。

平稳随机过程的严格定义是：若随机过程$X(t)$的所有有穷的n维分布函数或概率密度，对所有的t加上同一个τ值保持不变，即下式成立

$$F_n(x_1,\cdots\cdots,x_n;t_1,\cdots\cdots,t_n)=F_n(x_1,\cdots\cdots,x_n;t_1+\tau,\cdots\cdots,t_n+\tau) \tag{1-3-1a}$$

或

$$f_n(x_1,\cdots\cdots,x_n;t_1,\cdots\cdots,t_n)=f_n(x_1,\cdots\cdots,x_n;t_1+\tau,\cdots\cdots,t_n+\tau), \tag{1-3-1b}$$

则称$X(t)$为严平稳随机过程或狭义平稳随机过程。

根据平稳随机过程的定义式（1-3-1b）容易看到，平稳随机过程$X(t)$的一维概率密度与时间t无关，即

$$f_1(x,t)=f_1(x,t+\tau)=f_1(x), \tag{1-3-2a}$$

且它的二维概率密度只与t_1、t_2的时间间隔$\tau=t_2-t_1$有关，而与时间起点无关，即

$$f_2(x_1,x_2;t_1,t_2)=f_2(x_1,x_2,\tau) \tag{1-3-2b}$$

因此，严平稳随机过程$X(t)$的数学期望和方差为

$$E(X(t))=\mu_X(t)=\int_{-\infty}^{\infty}xf_1(x)dx=\mu_X \tag{1-3-3a}$$

$$\begin{aligned}D_X(t)&=\int_{-\infty}^{\infty}(x(t)-\mu_X(t))^2f_1(x,t)\,dx\\&=\int_{-\infty}^{\infty}(x-\mu_X)^2f_1(x)dx=D_X=\sigma_X^2\end{aligned} \tag{1-3-3b}$$

即严平稳随机过程的数学期望和方差均为常数。同时，$X(t)$的自协方差函数和自相关函数为

$$\begin{aligned}D_X(t_1,t_2)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{x(t_1)-\mu_X(t_1)\}\{x(t_2)\\&\quad-\mu_X(t_2)\}f_2(x_1,x_2;t_1,t_2)dx_1dx_2\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-\mu_X)(x_2-\mu_X)f_2(x_1,x_2;\tau)\\&\quad\cdot dx_1dx_2=D_X(\tau),\end{aligned} \tag{1-3-3c}$$