

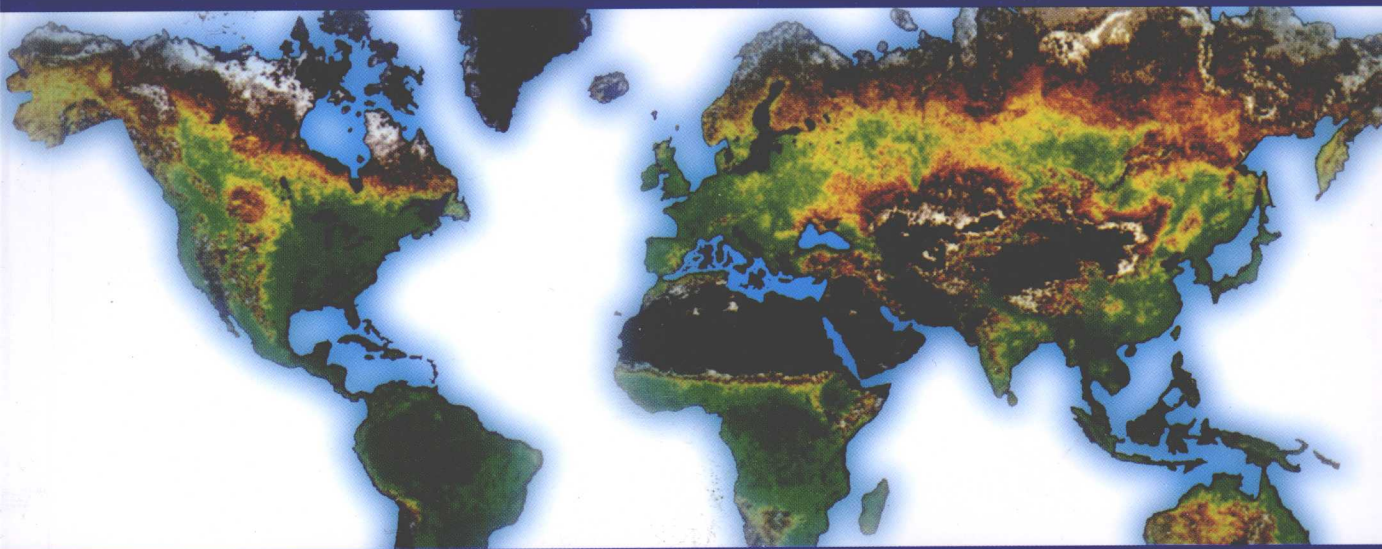
GEO-SPATIAL INFORMATION SCIENCE

● 普通高等教育测绘类规划教材

(第二版)

广义测量平差

崔希璋 於宗俦 陶本藻 刘大杰 编著
于正林 孙海燕 王新洲



WUHAN UNIVERSITY PRESS

武汉大学出版社

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 编著
于正林 孙海燕 王新洲



WUHAN UNIVERSITY PRESS

武汉大学出版社

图书在版编目(CIP)数据

广义测量平差/崔希璋等编著. —2版. —武汉:武汉大学出版社,2009.9
普通高等教育测绘类规划教材
ISBN 978-7-307-07268-8

I. 广… II. 崔…[等] III. 测量平差—高等学校—教材 IV. P207

中国版本图书馆 CIP 数据核字(2009)第 141601 号

责任编辑:王金龙 责任校对:王 建 版式设计:马 佳

出版发行:武汉大学出版社 (430072 武昌 珞珈山)

(电子邮件:cbs22@whu.edu.cn 网址:www.wdp.com.cn)

印刷:安陆市鼎鑫印务有限责任公司

开本:787×1092 1/16 印张:14 字数:352千字

版次:2001年3月第1版 2009年9月第2版

2009年第2版第1次印刷

ISBN 978-7-307-07268-8/P·161

定价:27.00元

版权所有,不得翻印;凡购买我社的图书,如有缺页、倒页、脱页等质量问题,请与当地图书销售部门联系调换。

第二版前言

《广义测量平差》作为测绘专业研究生教材已近 30 年了,1982 年 6 月由崔希璋教授主持,於宗俦、陶本藻、刘大杰编著的该教材首次出版,很快成为全国测绘类研究生的通用教材,获得有关教师和广大读者的高度好评。随着测绘技术和测量平差理论和方法的进步,1992 年由刘大杰、于正林主编,於宗俦、陶本藻参编的第二版教材,对第一版做了较大增补,列入了我们许多科研成果,极大地丰富了研究生的教学内容和该书的科学参考价值。2001 年根据研究生教学大纲要求和贯彻精选内容的少而精原则,由陶本藻主编对第二版做了修订和压缩,并由孙海燕、王新洲分别增写了该版的第 5、6 两章,出版了《广义测量平差》新版(即第三版)。根据多年来有关教师、研究生和读者的意见,特别是对本书贡献最大的作者刘大杰教授生前的修改建议,考虑教学用书的需要,在武汉大学出版社大力支持下,决定修订再版本书,由陶本藻教授主编修订。

在现行的各校测绘类研究生课程中,虽然课程名称不同,但有关测量数据处理总是一门必修课程,大部分院校仍采用这本《广义测量平差》教材,考虑到本教材的传统性、连续性和共知性,以及老一辈测绘专家的努力和贡献,本书没有更改书名和作者署名,本版实为《广义测量平差》第四版。

本版仍维持新版的教学体系,教学重点仍是前四章,基本内容不变。所修改和增补的主要是第 2、4 两章。主要是对其中某些内容进行扩充并深入了其理论解释,增加了方法上的应用和算例等。新增加的 2-10 节(向量空间理论中的平差问题)和 4-8 节至 4-11 节四节的卡尔曼滤波基础理论是刘大杰教授的遗作,是研究生学习有关内容所必须掌握的知识。

本书最后所列出的参考文献,是我们收集到的 5 本与课程有关的国内外出版的专著和教科书,是专门为广大研究生和读者进一步学习、研究和发展广义测量平差而准备的。

本书的修订和出版得到武汉大学出版社的大力支持,由于他们的辛勤和认真的工作,才能在很短时间内出版本书(第二版),以满足广大研究生和读者的需要,在此表示深切感谢。

希望能继续得到读者支持,不断提供改进意见。

编著者

2009 年 5 月于武汉

前 言

《广义测量平差》作为测绘专业研究生教材已近 20 年,1982 年 6 月第一次出版,1992 年 10 月经较大修改出版了第二版,至今仍作为许多高等院校测绘专业的研究生教材或参考书以及本科生高年级选修教材。根据教师和读者对教材内容的修改意见,考虑教学用书的需要,决定修订再版本书。

在一些学校的研究生教学计划中,将原开设的《广义测量平差》课程改名为《现代测量数据处理理论》,以原来的广义测量平差基本内容为主并进行某些扩充,本书的修订就考虑了这一变化因素。为了本书的连续性和共知性,我们没有更改书名,但书名后用括号加注“新版”字样,其内容可认为是《现代测量数据处理理论》课程的基本教材。

本版对第二版在教学体系上做了较大修改,并以少而精原则精选内容。修订主要之处是:将第二版的第 3 章秩亏自由网平差、第 5 章滤波与配置及第 6 章若干平差问题的补充等这三章所精选的教学内容合并为一章,取名为最小二乘平差的统一理论和方法;删除了第二版中第 1 章随机过程基础和附录等纯数学内容,删除了第二版的第 8 章卡尔曼滤波的渐近性质和误差分析;基本保留了第 2 章估计方法和广义测量平差原理、第 4 章平差随机模型的验后估计、第 7 章动态线性系统的卡尔曼滤波等三章内容;新增了稳健估计的基本理论和有偏估计两章。经上述修订,已将第二版中 8 章 53 节,缩成为 6 章 36 节,字数也大为减少。但仍保留了介绍近代测量平差理论和方法最新成果为主线的特色。

本书的第一版,由崔希璋主持,於宗俦、陶本藻、刘大杰参与编写;第二版由刘大杰、于正林主编,负责全面修订,於宗俦、陶本藻参编;本版由陶本藻主编,负责全面修订,孙海燕、王新洲参编,孙海燕编写了第 5 章,王新洲编写了第 6 章。刘经南参与了本书编写大纲的讨论,所提出的修订意见已反映在本版中。

《广义测量平差原理》第一版、第二版均由测绘出版社出版,经协商同意,本书改由武汉测绘科技大学出版社出版,我们对这两个出版社的领导和编辑的大力支持表示衷心感谢。

武汉测绘科技大学教务处和全国测绘类高等学校教学指导委员会对本书出版做了大量工作,我们在此也一并表示深切感谢。

希望能继续得到读者支持,不断提供改进意见。

编著者

2000 年 6 月于武汉

目 录

第二版前言	(1)
前言	(1)
第 1 章 估计方法和广义测量平差原理	(1)
1-1 概述	(1)
1-2 多维正态分布	(3)
1-3 极大似然估计	(9)
1-4 最小二乘估计	(16)
1-5 极大验后估计	(19)
1-6 最小方差估计	(21)
1-7 线性最小方差估计	(23)
1-8 贝叶斯估计	(25)
1-9 广义测量平差原理	(27)
第 2 章 最小二乘平差的统一理论和方法	(31)
2-1 概述	(31)
2-2 秩亏自由网平差	(37)
2-3 附加系统参数的自由网平差	(46)
2-4 极大验后滤波与推估	(56)
2-5 最小二乘配置	(64)
2-6 静态逐次滤波	(78)
2-7 随机模型具有奇异协因数阵的平差	(84)
2-8 广义 G-M 模型的平差问题	(87)
2-9 广义 G-M 模型下的精度和统计性质	(91)
2-10 向量空间理论中的平差问题	(95)
第 3 章 平差随机模型的验后估计	(102)
3-1 概述	(102)
3-2 赫尔默特方差估计法	(103)
3-3 方差-协方差分量估计	(114)
3-4 二次无偏估计法	(120)
3-5 方差分量估计中的精度评定	(129)
第 4 章 动态线性系统的卡尔曼滤波	(136)

4-1	连续线性系统的数学模型	(136)
4-2	离散线性系统的数学模型	(141)
4-3	离散线性系统的卡尔曼滤波	(145)
4-4	动态测量系统的卡尔曼滤波	(151)
4-5	离散型卡尔曼滤波的推广	(155)
4-6	离散线性系统的预测	(159)
4-7	离散线性系统的平滑	(163)
4-8	线性确定系统的能观性和能控性	(169)
4-9	卡尔曼滤波的稳定性	(173)
4-10	模型误差分析	(178)
4-11	滤波的发散现象和克服发散的方法	(181)
第 5 章	稳健估计的基本理论	(186)
5-1	统计稳健性	(186)
5-2	稳健性的数学描述	(192)
5-3	位置参数的稳健估计	(201)
第 6 章	有偏估计	(207)
6-1	概述	(207)
6-2	岭估计	(209)
6-3	广义岭估计	(213)
参考文献		(215)

Contents

The second edition preface	(1)
Preface	(1)
Chapter 1 Estimate method and the principle of generalized surveying adjustment ...	(1)
1-1 Introduction	(1)
1-2 Multinormal distribution	(3)
1-3 Maximum likelihood estimation	(9)
1-4 Least squares estimation	(16)
1-5 Maximum a posteriori estimation	(19)
1-6 Minimum variance estimation	(21)
1-7 Linear minimum variance estimation	(23)
1-8 Bayes estimation	(25)
1-9 The principle of generalized surveying adjustment	(27)
Chapter 2 The unified theory and method of least squares adjustment	(31)
2-1 Introduction	(31)
2-2 Adjustment of rank defect free network	(37)
2-3 Free network adjustment with systematic parameters	(46)
2-4 Maximum a posteriori filtering and prediction	(56)
2-5 Least squares collocation	(64)
2-6 Sequential static filtering	(78)
2-7 Adjustment with singular cofactor matrix in random model	(84)
2-8 Adjustment subjects of generalized G-M model	(87)
2-9 The precision and statistic property of generalized G-M model	(91)
2-10 Adjustment Problems in the Theory of Vector Space	(95)
Chapter 3 A posteriori estimation of the adjustment random model	(102)
3-1 Introduction	(102)
3-2 Helmert method of variance components estimation	(103)
3-3 Estimation of variance and covariance components	(114)
3-4 Method of quadratic unbiased estimation	(120)
3-5 Precision evaluating in variance components estimation	(129)

Chapter 4	Kalman filtering of dynamic linear system	(136)
4-1	The mathematical model of continuous linear system	(136)
4-2	The mathematical model of discrete linear system	(141)
4-3	Kalman filtering of discrete linear system	(145)
4-4	Kalman filtering of dynamic surveying system	(151)
4-5	Generalization of Kalman filtering of discrete system	(155)
4-6	The prediction of discrete linear system	(159)
4-7	The smoothing of discrete linear system	(163)
4-8	The Observability and Controllability of Determined Linear Systems	(169)
4-9	The Stability of Kalman Filtering	(173)
4-10	Model Error Analysis	(178)
4-11	Divergence Phenomenon of Filtering and Ways to Overcome the Divergence	(181)
Chapter 5	The basic theory of robust estimation	(186)
5-1	Statistic robustness	(186)
5-2	mathematical description of robustness	(192)
5-3	Robust estimation of location parameters	(201)
Chapter 6	Biased estimation	(207)
6-1	Introduction	(207)
6-2	Ridge estimation	(209)
6-3	Generalized ridge estimation	(213)
Reference books		(215)

第1章 估计方法和广义测量平差原理

1-1 概 述

在测量、通信和控制等学科中,为了求得某些未知参数,常常要进行一系列的观测.由于测量上的局限性,往往只能观测未知量的某些函数,且观测值中必然含有误差(或称为噪声).这就产生了根据含有误差的观测值求定未知参数估值的问题.下面举几个例子.

(1)为了确定平面或三维控制网中各点的坐标,对控制网的边长和方向(或坐标差)进行了观测,当然,观测值包含有误差.设各点的坐标为未知参数向量 X ,而包括边长和方向的观测值向量为 L ,则 L 和 X 之间有函数关系

$$L = F(X) + \Delta$$

式中 Δ 表示误差向量.通过含有误差 Δ 的观测向量 L 来求定待定点坐标的最佳估值,就是一个估计问题.在测量中,就是一个平差问题.

(2)通信理论中的一个重要问题是从接收到的信号中,提取被发送的信号.设被发送的信息调制成信号 $S(t)$,而接收到的信号也就是信号的观测值 $L(t)$,由于大气噪声和电路噪声的干扰,因此有

$$L(t) = S(t) + n(t)$$

其中 $n(t)$ 是噪声, t 表示时间.通信中的主要问题就是从 $L(t)$ 中将有用的信号 $S(t)$ 分离出来,也就是由 $L(t)$ 求定 $S(t)$ 的最佳估值.信号 $S(t)$ 也是一种未知参数.

(3)生产过程的自动化可以达到高效率和高精度.在实现生产过程的控制中,需要通过生产系统进行状态的不断测量,得到与系统运行状态有关的观测值;然后对观测值进行分析处理得到控制信号,实时地控制生产系统按要求运行.但由于观测值中存在误差,所以,为了得到控制信号,就要求由观测值来估计系统的运行状态.

(4)卫星(或其他运动体)的轨道往往可以由如下微分方程确定

$$\dot{X}(t) = f(X(t), U(t), \Omega(t))$$

式中 t 表示时间; $X(t)$ 表示卫星的轨道参数,在此处称为状态向量; $U(t)$ 为控制向量; $\Omega(t)$ 是随机的状态噪声.为了精确估计或预测卫星的轨道,就需要对卫星进行观测,从而得到大量的观测数据 $L(t)$,然后实时地由含有误差的观测值 $L(t)$ 来估计卫星的轨道,即估计卫星的轨道参数.

以上例中所述的信号或状态都可以说是一种未知参数.在测量平差中,通常称非随机的未知参数向量为参数,而称随机参数向量为信号,而称随时间 t 变化的动态系统中的未知参数向量为状态向量,或简称为状态.可以看到,在上面的例子中,都存在一个对未知参数进行估计的问题.

一般说来,若设 x 为 t 阶未知参数向量(简称为参数), L 为 n 阶观测向量(或称观测值), Δ

表示 n 维误差(或噪声)向量. 那么, 所谓估计问题, 就是根据含有误差 Δ 的观测值 L , 构造一个函数 $\hat{X}(L)$, 使 $\hat{X}(L)$ 成为未知参数向量 X 的最佳估计量, 其具体数值称为最佳估值(以后一般不区分其含义). 通常将 $\hat{X}(L)$ 简记为 $\hat{X}$, 并记

$$\Delta_{\hat{X}} = X - \hat{X}(L) = X - \hat{X}$$

称 $\Delta_{\hat{X}}$ 为 $\hat{X}(L)$ 的估计误差.

可以看到, 当 $\Delta_{\hat{X}}$ 的数学期望等于零时, $\Delta_{\hat{X}}$ 的方差就等于 $E(\Delta_{\hat{X}}\Delta_{\hat{X}}^T)$; 而当 X 为非随机量时, 未知参数的估值 $\hat{X}$ 的方差 $D_{\hat{X}}$ 也就等于其误差方差 $D(\Delta_{\hat{X}})$. 在估计理论中, 通常是用估计量 $\hat{X}$ 的误差方差 $D(\Delta_{\hat{X}})$ 来衡量其精度的. 但在经典的最小二乘平差中, 由于 X 一般都是非随机参数, 所以习惯上都用估值(平差值)的方差衡量精度.

在根据观测值 L 求未知参数 X 的估值 $\hat{X}(L)$ 时, 总是希望所得到的估值是最优的. 由估计理论知道, 最优估计量主要应具有以下几个性质:

(1) 一致性. 由观测值得到的估值 $\hat{X}(L)$ 通常与其真值是不同的, 我们希望当观测值个数 n 增加时, 估计量变得更好些; 当 n 无限增大时, 估计量向被估计的参数趋近的概率等于 1. 即如果对于任意 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P(X - \varepsilon < \hat{X} < X + \varepsilon) = 1 \quad (1-1-1)$$

则称估计量 $\hat{X}$ 具有一致性; 若有

$$\lim_{n \rightarrow \infty} ((X - \hat{X})(X - \hat{X})^T) = 0 \quad (1-1-2)$$

则称此估计量是均方一致的. 估计量的一致性是从它的极限性质来看的.

(2) 无偏性. 若估计量 $\hat{X}$ 的数学期望等于被估计量 X 的数学期望, 即

$$E(\hat{X}) = E(X) \quad (1-1-3)$$

如果 X 是非随机量, 上式即为

$$E(\hat{X}) = X \quad (1-1-4)$$

则称 $\hat{X}$ 为无偏估计量. 如果 $E(\hat{X}) \rightarrow X (n \rightarrow \infty)$, 则称 $\hat{X}$ 为渐近无偏.

(3) 有效性. 若由观测向量 L 得到无偏估计量 $\hat{X}$ 的误差方差 $E((X - \hat{X})(X - \hat{X})^T)$, 小于由 L 得到的任何其他无偏估计量 X^* 的误差方差 $E((X - X^*)(X - X^*)^T)$, 即

$$E((X - \hat{X})(X - \hat{X})^T) < E((X - X^*)(X - X^*)^T)$$

或写为

$$D(\Delta_{\hat{X}}) < D(\Delta_{X^*}) \quad (1-1-5)$$

则称 $\hat{X}$ 是有效估计量, 也称 $\hat{X}$ 具有有效性或方差最小性.

以不同的准则来求定未知参数的最佳估值, 可得到不同的估计方法. 估计方法主要有极大似然估计, 最小二乘估计, 极大验后估计, 最小方差估计和线性最小方差估计等; 经典的测量平差法都是以最小二乘估计或极大似然估计为根据导出的; 而滤波、配置和动态系统的卡尔曼滤波等, 最初是以极大验后估计或最小方差估计为根据导出的. 因此, 概率统计中的估计理论是广义测量平差的理论基础.

1-2 多维正态分布

正态分布是测量平差理论中最常用的误差分布,是最小二乘平差误差理论的基础. 本节在已学过的一元正态分布的基础上,对多维正态分布做全面阐述. 广义测量平差理论中还涉及其他分布,则将分别在相应章节中一一介绍.

1. 多维正态分布的定义和性质

已知随机变量 X 的正态分布概率密度为

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu_x)^2 \right\} \quad (1-2-1)$$

式中两个参数 μ_x 和 σ^2 分别为随机变量 X 的数学期望和方差. 当 $\mu_x = 0, \sigma^2 = 1$ 时, X 为标准正态分布变量, 记为 $X \sim N(0, 1)$, 其概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} x^2 \right\} \quad (1-2-2)$$

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

$$X_{n \times 1} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = A_{n \times m} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_m \end{pmatrix} + A_0_{n \times 1}$$

为 n 维正态随机向量. 此时, X 的数学期望和方差阵为

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

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

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

则

$$X_1 \sim N_r(\mu_1, D_{11}), X_2 \sim N_{n-r}(\mu_2, D_{22}).$$

2. 多维正态分布

设有 n 维正态随机向量 $X \sim N_n(\mu_x, D_x)$, 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

$$f(x) = 2(\pi)^{-\frac{n}{2}} |D_x|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x - \mu_x)^T D_x^{-1}(x - \mu_x)\right\} \quad (1-2-3)$$

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(1-2-3)式可得其概率密度为

$$f(x, y) = \frac{1}{2\pi \sqrt{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2}} \cdot \exp\left\{-\frac{(x - \mu_x)^2 \sigma_y^2 - 2(x - \mu_x)(y - \mu_y)\sigma_{xy} + (y - \mu_y)^2 \sigma_x^2}{2(\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2)}\right\}$$

因相关系数 $\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$, 所以上式可写为

$$f(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} \exp\left\{-\frac{1}{2(1 - \rho_{xy}^2)} \left[\frac{(x - \mu_x)^2}{\sigma_x^2} - 2\rho \frac{(x - \mu_x)(y - \mu_y)}{\sigma_x \sigma_y} + \frac{(y - \mu_y)^2}{\sigma_y^2} \right]\right\} \quad (1-2-4)$$

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

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi \sigma_x \sigma_y} \exp\left\{-\frac{(x - \mu_x)^2}{2\sigma_x^2} - \frac{(y - \mu_y)^2}{2\sigma_y^2}\right\} \\ &= \frac{1}{\sqrt{2\pi} \sigma_x} \exp\left\{-\frac{(x - \mu_x)^2}{2\sigma_x^2}\right\} \cdot \frac{1}{\sqrt{2\pi} \sigma_y} \exp\left\{-\frac{(y - \mu_y)^2}{2\sigma_y^2}\right\} \\ &= f_x(x) f_y(y) \end{aligned} \quad (1-2-5)$$

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(1-2-4)式绘制二维正态曲面(密度曲面)如图 1-1 所示。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。

3. 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

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

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

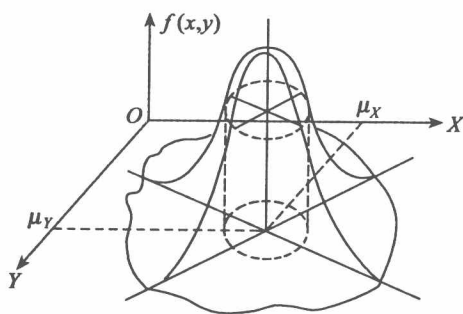


图 1-1

$$f(x) = (2\pi)^{-\frac{n+1}{2}} |D_x|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}^T D_x^{-1} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix} \right\} \quad (1-2-6)$$

按分块矩阵求逆公式,有

$$D_x^{-1} = \begin{bmatrix} D_{11}^{-1} + D_{11}^{-1} D_{12} \tilde{D}_{22}^{-1} D_{21} D_{11}^{-1} & -D_{11}^{-1} D_{12} \tilde{D}_{22}^{-1} \\ -\tilde{D}_{22}^{-1} D_{21} D_{11}^{-1} & \tilde{D}_{22}^{-1} \end{bmatrix} \quad (1-2-7)$$

或为

$$D_x^{-1} = \begin{bmatrix} \tilde{D}_{11}^{-1} & -\tilde{D}_{11}^{-1} D_{12} D_{22}^{-1} \\ -D_{22}^{-1} D_{21} \tilde{D}_{11}^{-1} & D_{22}^{-1} + D_{22}^{-1} D_{21} \tilde{D}_{11}^{-1} D_{12} D_{22}^{-1} \end{bmatrix} \quad (1-2-8)$$

其中

$$\left. \begin{aligned} \tilde{D}_{11} &= D_{11} - D_{12} D_{22}^{-1} D_{21} \\ \tilde{D}_{22} &= D_{22} - D_{21} D_{11}^{-1} D_{12} \end{aligned} \right\} \quad (1-2-9)$$

可将(1-2-7)和(1-2-8)两式分别写为

$$D_x^{-1} = \begin{bmatrix} -D_{11}^{-1} D_{12} \\ E_t \end{bmatrix} \tilde{D}_{22}^{-1} \begin{bmatrix} -D_{21} D_{11}^{-1} & E_t \end{bmatrix} + \begin{bmatrix} D_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \quad (1-2-10)$$

$$D_x^{-1} = \begin{bmatrix} E_n \\ -D_{22}^{-1} D_{21} \end{bmatrix} \tilde{D}_{11}^{-1} \begin{bmatrix} E_n & -D_{12} D_{22}^{-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & D_{22}^{-1} \end{bmatrix} \quad (1-2-11)$$

因 D_x 还可分解为

$$D_x = \begin{bmatrix} D_{11} & 0 \\ D_{21} & \tilde{D}_{22} \end{bmatrix} \begin{bmatrix} E & D_{11}^{-1} D_{12} \\ 0 & E \end{bmatrix} = \begin{bmatrix} \tilde{D}_{11} & D_{12} \\ 0 & D_{22} \end{bmatrix} \begin{bmatrix} E & 0 \\ D_{22}^{-1} D_{21} & E \end{bmatrix} \quad (1-2-12)$$

所以, D_x 的行列式之值为

$$|D_x| = |D_{11}| |\tilde{D}_{22}| = |D_{22}| |\tilde{D}_{11}| \quad (1-2-13)$$

利用(1-2-10)、(1-2-9)式和(1-2-13)式,可将概率密度(1-2-6)式改写为

$$\begin{aligned} f(x) &= f(x_1, x_2) \\ &= (2\pi)^{-\frac{n}{2}} |D_{11}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x_1 - \mu_1)^T D_{11}^{-1}(x_1 - \mu_1)\right\} \cdot \\ &\quad (2\pi)^{-\frac{1}{2}} |\tilde{D}_{22}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x_2 - \tilde{\mu}_2)^T \tilde{D}_{22}^{-1}(x_2 - \tilde{\mu}_2)\right\} \end{aligned} \quad (1-2-14)$$

或

$$\begin{aligned} f(x) &= f(x_1, x_2) \\ &= (2\pi)^{-\frac{1}{2}} |D_{22}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x_2 - \mu_2)^T D_{22}^{-1}(x_2 - \mu_2)\right\} \cdot \\ &\quad (2\pi)^{-\frac{n}{2}} |\tilde{D}_{11}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x_1 - \tilde{\mu}_1)^T \tilde{D}_{11}^{-1}(x_1 - \tilde{\mu}_1)\right\} \end{aligned} \quad (1-2-15)$$

其中

$$\begin{aligned} \tilde{\mu}_1 &= \mu_1 + D_{12}D_{22}^{-1}(x_2 - \mu_2) \\ \tilde{\mu}_2 &= \mu_2 + D_{21}D_{11}^{-1}(x_1 - \mu_1) \end{aligned} \quad (1-2-16)$$

根据边际概率密度和多维正态分布的性质可知

$$f_1(x_1) = (2\pi)^{-\frac{n}{2}} |D_{11}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x_1 - \mu_1)^T D_{11}^{-1}(x_1 - \mu_1)\right\} \quad (1-2-17)$$

$$f_2(x_2) = (2\pi)^{-\frac{1}{2}} |D_{22}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x_2 - \mu_2)^T D_{22}^{-1}(x_2 - \mu_2)\right\} \quad (1-2-18)$$

又由条件概率密度公式知

$$f(x_2/x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} \quad (1-2-19)$$

$$f(x_1/x_2) = \frac{f(x_1, x_2)}{f_2(x_2)} \quad (1-2-20)$$

将(1-2-14)和(1-2-17)两式代入(1-2-19)式,得

$$f(x_2/x_1) = (2\pi)^{-\frac{1}{2}} |\tilde{D}_{22}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x_2 - \tilde{\mu}_2)^T \tilde{D}_{22}^{-1}(x_2 - \tilde{\mu}_2)\right\} \quad (1-2-21)$$

而将(1-2-15)和(1-2-18)两式代入(1-2-20)式,即得

$$f(x_1/x_2) = (2\pi)^{-\frac{n}{2}} |\tilde{D}_{11}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x_1 - \tilde{\mu}_1)^T \tilde{D}_{11}^{-1}(x_1 - \tilde{\mu}_1)\right\} \quad (1-2-22)$$

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

$$\begin{aligned} E(X_1/x_2) &= \tilde{\mu}_1 = \mu_1 + D_{12}D_{22}^{-1}(x_2 - \mu_2) \\ E(X_2/x_1) &= \tilde{\mu}_2 = \mu_2 + D_{21}D_{11}^{-1}(x_1 - \mu_1) \end{aligned} \quad (1-2-23)$$

$$\begin{aligned} D(X_1/x_2) &= \tilde{D}_{11} = D_{11} - D_{12}D_{22}^{-1}D_{21} \\ D(X_2/x_1) &= \tilde{D}_{22} = D_{22} - D_{21}D_{11}^{-1}D_{12} \end{aligned} \quad (1-2-24)$$

因此,(1-2-21)和(1-2-22)式又可写为

$$\left. \begin{aligned} f(x_2/x_1) &= (2\pi)^{-\frac{1}{2}} |D(X_2/x_1)|^{-\frac{1}{2}} \cdot \\ &\exp\left\{-\frac{1}{2}[x_2 - E(X_2/x_1)]^T \cdot \right. \\ &\quad \left. D^{-1}(X_2/x_1)[x_2 - E(X_2/x_1)]\right\} \\ f(x_1/x_2) &= (2\pi)^{-\frac{1}{2}} |D(X_1/x_2)|^{-\frac{1}{2}} \cdot \\ &\exp\left\{-\frac{1}{2}[x_1 - E(X_1/x_2)]^T \cdot \right. \\ &\quad \left. D^{-1}(X_1/x_2)[x_1 - E(X_1/x_2)]\right\} \end{aligned} \right\} \quad (1-2-25)$$

正态分布的条件期望具有以下性质:

(1) 由(1-2-23)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

$$\left. \begin{aligned} \tilde{X} &= X - E(X/Y) \\ Z &= AY \end{aligned} \right\} \quad (1-2-26)$$

则 $\tilde{X}$ 是与 Z 互相独立的随机向量. 这是因为

$$\begin{aligned} \tilde{X} &= X - \{\mu_X + D_{XY}D_Y^{-1}(Y - \mu_Y)\} \\ &= X - D_{XY}D_Y^{-1}Y + D_{XY}D_Y^{-1}\mu_Y \end{aligned}$$

由协方差传播律可得

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

(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 $\text{cov}(Y_1, Y_2) = 0$, 而 $D(X, Y_1) = D_{XY_1} \neq 0$, $D(X, Y_2) = D_{XY_2} \neq 0$, 则有

$$E(X/Y) = E(X/Y_1, Y_2) = E(X/Y_1) + E(X/Y_2) - \mu_X \quad (1-2-27)$$

证 因为

$$\begin{aligned} E(X/Y) &= E(X/Y_1, Y_2) = \mu_X + D_{XY}D_Y^{-1}(Y - \mu_Y) \\ &= \mu_X + [D_{XY_1} \quad D_{XY_2}] \begin{bmatrix} D_1^{-1} & 0 \\ 0 & D_2^{-1} \end{bmatrix} \begin{bmatrix} Y_1 - \mu_1 \\ Y_2 - \mu_2 \end{bmatrix} \end{aligned}$$

所以

$$\begin{aligned} E(X/Y_1, Y_2) &= \{\mu_X + D_{XY_1}D_1^{-1}(Y_1 - \mu_1)\} + \\ &\quad \{D_{XY_2}D_2^{-1}(Y_2 - \mu_2) + \mu_X - \mu_X\} \\ &= E(X/Y_1) + E(X/Y_2) - \mu_X \end{aligned}$$

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

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

令 $\tilde{Y}_2 = Y_2 - E(Y_2/Y_1)$, 则有

$$\begin{aligned}
 E(X/y_1, y_2) &= E(X/y_1, \tilde{y}_2) \\
 &= E(X/y_1) + E(X/\tilde{y}_2) - \mu_x
 \end{aligned} \tag{1-2-28}$$

证 因为

$$\begin{aligned}
 \tilde{Y}_2 &= Y_2 - \mu_2 - D_{21}D_{11}^{-1}(Y_1 - \mu_1) \\
 &= [-D_{21}D_{11}^{-1} \quad E] \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} - \mu_2 + D_{21}D_{11}^{-1}\mu_1
 \end{aligned}$$

所以

$$\left. \begin{aligned}
 E(\tilde{Y}_2) &= 0, \\
 D(\tilde{Y}_2) &= [-D_{21}D_{11}^{-1} \quad E] \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} -D_{11}^{-1}D_{12} \\ E \end{bmatrix} \\
 &= D_{22} - D_{21}D_{11}^{-1}D_{12} = \tilde{D}_{22} \\
 D(\tilde{Y}_2, Y_1) &= 0, \\
 D(\tilde{Y}_2, X) &= [-D_{21}D_{11}^{-1} \quad E] \begin{bmatrix} D_{Y_1X} \\ D_{Y_2X} \end{bmatrix} \\
 &= D_{Y_2X} - D_{21}D_{11}^{-1}D_{Y_1X}
 \end{aligned} \right\} \tag{1-2-29}$$

利用分块求逆公式和(1-2-29)式得

$$\begin{aligned}
 E(X/y_1, y_2) &= \mu_x + D_{xY}D_Y^{-1}(Y - \mu_Y) \\
 &= \mu_x + [D_{xY_1} \quad D_{xY_2}] \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}^{-1} \begin{bmatrix} Y_1 - \mu_1 \\ Y_2 - \mu_2 \end{bmatrix} \\
 &= \mu_x + [D_{xY_1} \quad D_{xY_2}] \left\{ \begin{bmatrix} -D_{11}^{-1}D_{12} \\ E \end{bmatrix} \right\} \\
 &\quad \tilde{D}_{22}^{-1} [-D_{21}D_{11}^{-1} \quad E] + \begin{bmatrix} D_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Y_1 - \mu_1 \\ Y_2 - \mu_2 \end{bmatrix} \\
 &= \mu_x + \{ (-D_{xY_1}D_{11}^{-1}D_{12} + D_{xY_2})D^{-1} \cdot (\tilde{Y}_2) \cdot \\
 &\quad [-D_{21}D_{11}^{-1} \quad E] + [D_{xY_1}D_{11}^{-1} \quad 0] \} \begin{bmatrix} Y_1 - \mu_1 \\ Y_2 - \mu_2 \end{bmatrix} \\
 &= \mu_x + D(X, \tilde{Y}_2)D^{-1}(\tilde{Y}_2) \{ -D_{21}D_{11}^{-1}(Y_1 - \mu_1) + \\
 &\quad Y_2 - \mu_2 \} + D_{xY_1}D_{11}^{-1}(Y_1 - \mu_1) + \mu_x - \mu_x \\
 &= E(X/y_1) + E(X/\tilde{y}_2) - \mu_x
 \end{aligned}$$

4. 矩阵反演公式

由于正定矩阵的逆阵唯一,故由(1-2-7)、(1-2-8)两式直接可得:

$$(D_{11} - D_{12}D_{22}^{-1}D_{21})^{-1} = D_{11}^{-1} + D_{11}^{-1}D_{12}(D_{22} - D_{21}D_{11}^{-1}D_{12})^{-1}D_{21}D_{11}^{-1} \tag{1-2-30}$$