航空高等院校教材

最佳线性滤波

——维纳滤波、卡尔曼滤波与最小二乘滤波

郑政谋 编

航空专业教材编审组

最佳线性滤波

——维纳滤波、卡尔曼滤波与最小二乘滤波

郑政谋 编

航空专业教材编审组

内 容 简 介

本书以尽可能简明的形式，介绍最佳线性滤波的最基本原理和方法。全书共分七章，在第一、二章的基础上，主要讨论维纳滤波、卡尔曼滤波和最小二乘滤波等三种基本的最佳线性滤波方法，为控制与信息类型专业学习随机控制，随机信号处理等方面的专业课程，进一步阅读估计理论方面的文献资料，和初步应用滤波方法解决工程技术问题提供必要的理论基础。

本书是根据航空工业部高等院校教材出版规划编写的，可供控制与信息类型专业高年级学生和低年级研究生选用；也可供有关专业的工程技术人员参考。

主审稿人 万百五
责任编辑 严智民

最佳线性滤波

郑政谋 编

*

西北工业大学印刷厂印装 内部发行

*

787×1092 1/16 印张 11.5 282千字

1983年12月第一版 1983年12月第一次印刷 印数 0001—1500册

统一书号：16144 j 定价 1.20元

前　言

由于工程技术上精度的要求不断提高，如何最有效地从被随机噪声污染的信号中提取出我们所需要的有用信息，已成为当前许多工程技术问题中逼切需要解决的共同课题；于是，最佳线性滤波，特别是六十年代初出现的卡尔曼滤波，也就逐渐成为许多不同工科专业所共同需要的专业基础理论。为了适应控制与信息类型专业的需要，我们编写了这本教材，以尽可能简明的形式，介绍最佳线性滤波最基本的原理和方法，作为这些专业大学本科学生的选修教材，也可以作为低年级研究生的专业基础课教材。

本书是最佳线性滤波的一本入门书。全书共分七章，主要讨论现代线性滤波中两类最基本的滤波方法——维纳滤波与卡尔曼滤波；第一章，为使只具有工科大学概率论基础的读者能够顺利地阅读全书各章节内容，特将概率论中与滤波有关的一些主要内容作一简单的回顾、补充与推广；第二章，作为维纳滤波与卡尔曼滤波的公共理论基础，概括地介绍一些估计理论中的有关概念，使读者在阅读以后各章时，有一总的线索和工具；第三章，讨论现代线性滤波中较早的一种滤波方法——维纳滤波，这是四十年代至六十年代之间的一种最主要的线性滤波方法；第四章，作为维纳滤波的一个重要发展，通过随机变量投影法，讨论了离散型与连续型卡尔曼滤波的基本方程；第五章，为进一步扩大卡尔曼滤波的适用范围，讨论了卡尔曼滤波的一些推广问题，如预测问题、平滑问题、有色噪声问题、非线性模型问题与发散及其控制方法问题等；第六章，揭示出这两大类线性滤波——维纳滤波与卡尔曼滤波的内在等效关系；第七章，介绍一种适用于完全缺乏统计知识情况的线性滤波方法——最小二乘滤波，并分析了它与卡尔曼滤波在算法上的类似性。这些内容，为读者进一步阅读估计理论方面的文献资料，和初步应用滤波方法解决工程技术问题，提供了必要的基础。

本书只要求读者具备工科大学概率论与线性代数方面的数学基础，就可以循序渐进地阅读全书内容。第三、四章虽然接触到一些线性系统方面的概念和公式，但这些内容都已以预备知识的形式在 3.1 和 4.1 节中作了简单的介绍，并在附录D中给出了线性系统可控、可测性和滤波稳定性的一些有关定义和定理，以备阅读时查用；当然，如果读者事先具有自动控制方面的初步知识，阅读时会更顺利点。本科生可只选读目录中不带*号的章节；但对低年级研究生，作为一门滤波基础课程，可以学完全书的全部内容；其中维纳滤波是否要学，可根据专业需要决定。

本书承西安交通大学万百五、汪荣鑫、王月娟三位同志详细审阅了全书底稿，并提出了许多宝贵的改进意见，本校林树根同志为本书绘制了全部插图底稿，在此谨致深切谢意。

由于水平所限，书中缺点错误在所难免，恳请读者批评指正。

编者

1983年1月

目　录

第一章　预备知识——概率基础的简单回顾

第二章　估计的基本概念

第三章　维纳滤波

第四章 卡尔曼滤波的基本方程

第五章 离散型卡尔曼滤波的推广

第六章* 平稳情况下卡尔曼滤波与维纳滤波的比较

第七章 最小二乘滤波

附　　录

参　考　文　献

第一章 预备知识——概率基础的简单回顾

最佳线性滤波所要解决的问题，是如何从带有随机噪声的信号观测值中，处理（估计）出我们所需要的随机信息。由于在滤波计算中，被处理的信号观测值，与被估计的信息，都是事先不可确切知道的（即带有随机性的），于是在建立滤波方案时，就不可避免地要接触到概率统计的一些概念和方法。这一章在假设读者已具备工科概率论基础，但对概率计算中的矢量矩阵运算不熟悉，对随机过程知识不足的情况下，把概率论中与最佳线性滤波有关的一些主要内容，作一简单的回顾、补充和推广。

§1.1 随机变量及其概率分布

如果对于随机试验E中的每一基本事件ω，都给定一个实数$X(\omega)$和它对应，这样定义在随机试验E的全体基本事件（样本空间）Ω上的单值实函数$X(\omega)$就是一个随机变量，简记作X。随机变量根据可能值（可能出现的数值）情况的不同，有两种主要的类型：

1.1.1 离散型的随机变量

可能值可数的随机变量称为离散型随机变量。设X是一个离散型的随机变量，其可能值是$x_1,x_2,\cdots$，X取得这些可能值的概率分别是$p_1,p_2,\cdots$；由于这一组数值完整地给出了X取得各个可能值的概率，因此给定了概率数列$p_1,p_2,\cdots$，就完全描述了离散型随机变量X的概率分布，但在滤波问题中，我们需要了解的并不是X的全面概率分布情况如何，而是刻划它的两个主要分布特性的数字特征如下：

（1） 数学期望

如果随机变量X的可能值是$x_1,x_2,\cdots$，其对应的概率是$p_1,p_2,\cdots$，则称①

$$E\{X\}\triangleq\sum_{i=1}^{\infty}x_ip_i \tag{1.1-1}$$

为离散型随机变量X的数学期望或均值，X的数学期望有时又写为μ_x。从式（1.1-1）可以看出，数学期望μ_x是随机变量X的所有可能值$x_1,x_2,\cdots$，以其对应的概率$p_1,p_2,\cdots$为权的加权平均值，因此μ_x刻划了随机变量X在数轴上的平均出现位置。

（2） 方差

随机变量X对其数学期望μ_x偏差平方$(X-\mu_x)^2$的数学期望

$$Var(X)\triangleq E\{(X-\mu_x)^2\}=\sum_{i=1}^{\infty}(x_i-\mu_x)^2p_i \tag{1.1-2}$$

称为随机变量X的方差，X的方差有时又记作σ_x^2或V_x。从式（1.1-2）可以看出，方差是随机变量$(X-\mu_x)^2$的所有可能值$(x_i-\mu_x)^2(i=1,2,\cdots)$以其对应的概率$p_i(i=1,2,\cdots)$为

①符号"$\triangleq$"表示"定义为"。

权的加权平均值，因此它刻划了X对其均值（数学期望）的平均离散程度。

1.1.2 连续型随机变量

可能值连续充满在某一区间上的随机变量称为连续型随机变量。设X是一个随机变量，它的可能值X连续充满在某一区间上，比如说$-\infty<x<\infty$；如果有一函数$f(x)$，对于任一区间(a,b)，概率

$$P(a<X\leqslant b)=\int_a^b f(x)dx \tag{1.1-3}$$

则称$f(x)$为连续型随机变量X的分布密度（或概率密度函数），并称

$$F(x)\triangleq P(X\leqslant x)=\int_{-\infty}^{x} f(x)dx \tag{1.1-4}$$

为X的分布函数。根据式(1.1-4)，显然当$f(x)$在x点连续时，有

$$F'(x)=f(x) \tag{1.1-5}$$

式(1.1-3)给我们指出，知道了$f(x)$就可以给出X出现在任一区间(a,b)上的概率，因此连续型随机变量X的概率密度函数$f(x)$也完全描述了X的概率分布。由式(1.1-3)及式(1.1-4)还可以看出，概率密度函数$f(x)$具有以下两个基本性质：

(i) $\int_{-\infty}^{\infty} f(x)dx=1$

(ii) $f(x)\geqslant 0$

类似于离散型随机变量的数学期望与方差定义(1.1-1)及(1.1-2)，可以给出连续型随机变量X的数学期望$E\{X\}$与方差$Var(X)$的定义如下：

$$E\{X\}\triangleq\int_{-\infty}^{\infty} xf(x)dx \tag{1.1-6}$$

$$Var(X)\triangleq\int_{-\infty}^{\infty}(x-\mu_x)^2f(x)dx \tag{1.1-7}$$

数学期望与方差具有以下两个常用的性质：

(i) 如果C为非随机量，X为随机变量，则有

$$E\{CX\}=CE\{X\}$$
$$Var(CX)=C^2Var(X)$$

(ii) 对于n个任意的随机变量$X_1,X_2,\cdots,X_n$，

$$E\{X_1+X_2+\cdots+X_n\}=E\{X_1\}+E\{X_2\}+\cdots+E\{X_n\}$$

对于n个相互独立（“独立”的意义以后说明）的随机变量$X_1,X_2,\cdots,X_n$

$$Var(X_1+X_2+\cdots+X_n)=Var(X_1)+Var(X_2)+\cdots+Var(X_n)$$

§1.2 随机矢量(多维随机变量)及其概率分布

设有随机矢量$(X_1X_2\cdots X_n)^T$，其中$X_i(i=1,2,\cdots,n)$都是随机变量，如果有一n元函数$f(x_1,x_2,\cdots,x_n)$，对于任一n维空间域V_n，概率

$$P\{(X_1X_2\cdots X_n)^T\in V_n\}=\int\overset{(n)}{\cdots}\int_{V_n} f(x_1,x_2,\cdots,x_n)dx_1dx_2\cdots dx_n \tag{1.2-1}$$

则称 $f(x_1,x_2,\cdots,x_n)$ 为 $(X_1X_2\cdots X_n)^T$ 的联合分布密度（或联合概率密度函数），并称

$$F(x_1,x_2,\cdots x_n)\triangleq P(X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n)$$
$$=\int_{-\infty}^{x_1}\int_{-\infty}^{x}\cdots\int_{-\infty}^{x_n}f(x_1,x_2,\cdots,x_n)dx_1dx_2\cdots dx_n \qquad (1.2\text{-}2)$$

为 $(X_1X_2\cdots X_n)^T$ 的联合分布函数。根据式（1.2-2），显然，当 $f(x_1,x_2,\cdots,x_n)$ 在 $(x_1,x_2,\cdots,x_n)$ 点连续时，有

$$\frac{\partial^n}{\partial x_1\partial x_2\cdots\partial x_n}F(x_1,x_2,\cdots,x_n)=f(x_1,x_2,\cdots,x_n) \qquad (1.2\text{-}3)$$

由式（1.2-1）可以看出，知道了 $f(x_1,x_2,\cdots,x_n)$ 就能给出 $(X_1X_2\cdots X_n)^T$ 出现在任一 n 维空间域 V_n 中的概率，因此联合分布密度 $f(x_1,x_2,\cdots,x_n)$ 完全描述了随机矢量 $(X_2X_2\cdots X_n)^T$ 的概率分布。

对于随机矢量，除了联合分布外，还可以推引出另外两种概率分布——边沿分布与条件分布：

边沿分布密度

边沿分布密度是描述多维随机变量中的某一个或某几个分量的概率分布的概率密度函数。根据概率乘法定理与分布密度的定义，可以证明

（i） $(X_1X_2\cdots X_n)^T$ 关于 X_1 的边沿分布密度

$$f_1(x_1)=\int_{-\infty}^{\infty}\overset{(n-1)}{\cdots}\int_{-\infty}^{\infty}f(x_1,x_2,\cdots,x_n)dx_2dx_3\cdots dx_n \qquad (1.2\text{-}4)$$

（ii） $(X_1X_2\cdots X_n)^T$ 关于 X_2 的边沿分布密度

$$f_2(x_2)=\int_{-\infty}^{\infty}\overset{(n-1)}{\cdots}\int_{-\infty}^{\infty}f(x_1,x_2,\cdots,x_n)dx_1dx_3\cdots dx_n \qquad (1.2\text{-}5)$$

（iii） $(X_1X_2\cdots X_n)^T$ 关于 $(X_1X_2\cdots X_{n-1})^T$ 的边沿分布密度

$$f_{1,2,\ldots,n-1}(x_1,x_2,\cdots,x_{n-1})=\int_{-\infty}^{\infty}f(x_1,x_2,\cdots,x_n)dx_n \qquad (1.2\text{-}6)$$

（iv） $(X_1X_2\cdots X_nY_1Y_2\cdots Y_m)^T$ 关于 $(X_1X_2\cdots X_n)^T$ 的边沿分布密度

$$f_x(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{\infty}\overset{(m)}{\cdots}\int_{-\infty}^{\infty}f(x_1,x_2,\cdots,x_n;y_1,y_2,\cdots,y_m)dy_1dy_2\cdots dy_m \qquad (1.\text{-}7)$$

条件分布密度

条件分布密度是在已知一个随机矢量的某些分量的出现值的条件下，其余分量的分布密度。根据事件概率的基本概念，可以得出

（i） 在 $X_1=x_1$，$X_2=x_2$，$\cdots$，$X_{n-1}=x_{n-1}$ 的条件下，X_n 的条件分布密度

$$f(x_n|x_1,x_2,\cdots,x_{n-1})=\frac{f(x_1,x_2\cdots,x_n)}{f_{1,2,\ldots,n-1}(x_1,x_2,\cdots,x_{n-1})}$$
$$=\frac{f(x_1,x\cdots,x_{,n-1}|x_n)f_n(x_n)}{f_{1,2,\ldots,n-1}(x_1,x_2,\cdots,x_{n-1})} \qquad (1.2\text{-}8)$$

(ii) 在 $(Y_1Y_2\cdots Y_m)^T=(y_1y_2\cdots y_m)^T$ 的条件下，$(X_1X_2\cdots X_n)^T$ 的条件分布密度

$$f(x_1,x_2,\cdots,x_n|y_1,y_2,\cdots,y_m)=\frac{f(x_1,x_2,\cdots,x_n;y_1,y_2,\cdots,y_m)}{f_y(y_1,y_2,\cdots,y_m)}$$

$$=\frac{f(y_1,y_2,\cdots,y_m|x_1,x_2,\cdots,x_n)f_x(x_1,x_2,\cdots,x_n)}{f_y(y_1,y_2,\cdots,y_m)} \quad (1.2\text{-}9)$$

公式 (1.2-8) 及 (1.2-9) 又称为分布密度的贝叶斯 (Bayes) 公式，是估计理论中的两个重要关系式。

采用矢量矩阵表示，记

$$\boldsymbol{X}\triangleq(X_1X_2\cdots X_n)^T$$

$$\boldsymbol{x}\triangleq(x_1x_2\cdots x_n)^T$$

$$\boldsymbol{X}\leqslant\boldsymbol{x}\triangleq(X_1\leqslant x_1\ X_2\leqslant x_2\cdots X_n\leqslant x_n)^T$$

$$\boldsymbol{x}_{-i}\triangleq(x_1x_2\cdots x_{i-1}\ x_{i+1}\cdots x_n)^T$$

$$d\boldsymbol{x}\triangleq dx_1dx_2\cdots dx_n$$

$$\mathbf{d}\boldsymbol{x}\triangleq(dx_1dx_2\cdots dx_n)^T$$

$$f(\boldsymbol{x})\triangleq f(x_1,x_2,\cdots,x_n)$$

$$f(\boldsymbol{x}|\boldsymbol{y})\triangleq f(x_1,x_2,\cdots,x_n|y_1,y_2,\cdots,y_m)$$

公式 (1.2-1)～(1.2-9) 可以写成更简洁的形式如下：

(1.2-1)： $P(\boldsymbol{X}\in V_n)=\int_{V_n}f(\boldsymbol{x})d\boldsymbol{x}$

(1.2-2)： $F(\boldsymbol{x})=P(\boldsymbol{X}\leqslant\boldsymbol{x})=\int_{-\infty}^{\boldsymbol{x}}f(\boldsymbol{x})d\boldsymbol{x}$

(1.2-3)： $\frac{\partial^n}{\partial\boldsymbol{x}}F(\boldsymbol{x})=f(\boldsymbol{x})$ （注意这里 $\partial\boldsymbol{x}=\partial x_1\partial x_2\cdots\partial x_n$）

(1.2-4)： $f_1(x_1)=\int_{-\infty}^{\infty}f(\boldsymbol{x})d\boldsymbol{x}_{-1}$

(1.2-5)： $f_2(x_2)=\int_{-\infty}^{\infty}f(\boldsymbol{x})d\boldsymbol{x}_{-2}$

(1.2-6)： $f_{1,2,\cdots,n-1}(\boldsymbol{x}_{-n})=\int_{-\infty}^{\infty}f(\boldsymbol{x})dx_n$

(1.2-7)： $f_x(\boldsymbol{x})=\int_{-\infty}^{\infty}f(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{y}$

(1.2-8)： $f(x_n|\boldsymbol{x}_{-n})=\frac{f(\boldsymbol{x})}{f_{1,2,\cdots,n-1}(\boldsymbol{x}_{-n})}=\frac{f(\boldsymbol{x}_{-n}|x_n)f_n(x_n)}{f_{1,2,\cdots,n-1}(\boldsymbol{x}_{-n})}$

(1.2-9)： $f(\boldsymbol{x}|\boldsymbol{y})=\frac{f(\boldsymbol{x},\boldsymbol{y})}{f_y(\boldsymbol{y})}=\frac{f(\boldsymbol{y}|\boldsymbol{x})f_x(\boldsymbol{x})}{f_y(\boldsymbol{y})}$

与一维随机变量一样，在滤波技术中，最常用的还不是多维随机变量的各类分布密度，而是它的某些数字特征如下：

(1) 均值矢量

$$\boldsymbol{\mu}_x=E\{\boldsymbol{X}\}\triangleq E\begin{bmatrix}X_1\\X_2\\\vdots\\X_n\end{bmatrix}=\begin{bmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{bmatrix} \quad (1.2\text{-}10)$$

其中 $\mu_i = E\{X_i\}(i=1,2,\cdots,n)$。均值矢量 $\boldsymbol{\mu}_x$ 是随机矢量 $\boldsymbol{X}$ 的所有可能出现的矢量的一个平均（概率意义上的）矢量。

（2） 方差阵

$$Var(\boldsymbol{X}) \triangleq E\{(\boldsymbol{X}-\boldsymbol{\mu}_x)(\boldsymbol{X}-\boldsymbol{\mu}_x)^T\}$$

$$= E\begin{bmatrix} (X_1-\mu_1)^2 & (X_1-\mu_1)(X_2-\mu_2) & \cdots & (X_1-\mu_1)(X_n-\mu_n) \\ (X_2-\mu_2)(X_1-\mu_1) & (X_2-\mu_2)^2 & \cdots & (X_2-\mu_2)(X_n-\mu_n) \\ \vdots & \vdots & & \vdots \\ (X_n-\mu_n)(X_1-\mu_1) & (X_n-\mu_n)(X_2-\mu_2) & \cdots & (X_n-\mu_n)^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} \qquad (1.2\text{-}11)$$

其中 $\sigma_i^2 \triangleq E\{(X_i-\mu_i)^2\}$，$\sigma_{ij} \triangleq E\{(X_i-\mu_i)(X_j-\mu_j)\}$。

方差阵 $Var(\boldsymbol{X})$ 是随机矢量 $\boldsymbol{X}$ 的各分量 $X_i(i=1,2,\cdots,n)$ 对其均值 $\mu_i(i=1,2,\cdots,n)$ 偏差的两两乘积平均值的矩阵，主对角线上各元素 $\sigma_i^2(i=1,2,\cdots,n)$ 是 $\boldsymbol{X}$ 的各分量的方差，主对角线两侧的元素 $\sigma_{ij}=E\{(X_i-\mu_i)(X_j-\mu_j)\}(i\neq j)$ 是 X_i 与 X_j 的协方差，σ_{ij} 的值在一定程度上反映 X_i 与 X_f 之间线性相关性的强弱程度。

（3） $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的协方差阵

$$Cov(\boldsymbol{X},\boldsymbol{Y}) \triangleq E\{(\boldsymbol{X}-\boldsymbol{\mu}_x)(\boldsymbol{Y}-\boldsymbol{\mu}_y)^T\}$$

$$= E\begin{bmatrix} (X_1-\mu_{x_1})(Y_1-\mu_{y_1}) & (X_1-\mu_{x_1})(Y_2-\mu_{y_2}) & \cdots & (X_1-\mu_{x_1})(Y_m-\mu_{ym}) \\ (X_2-\mu_{x_2})(Y_1-\mu_{y_1}) & (X_2-\mu_{x2})(Y_2-\mu_{y_2}) & \cdots & (X_2-\mu_{x_2})(Y_m-\mu_{ym}) \\ \vdots & \vdots & & \vdots \\ (X_n-\mu_{xn})(Y_1-\mu_{y_1}) & (X_n-\mu_{xn})(Y_2-\mu_{y_2}) & \cdots & (X_n-\mu_{xn})(Y_m-\mu_{ym}) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{x_1y_1} & \sigma_{x_1y_2} & \cdots & \sigma_{x_1ym} \\ \sigma_{x_2y_1} & \sigma_{x_2y_2} & \cdots & \sigma_{x_2ym} \\ \vdots & \vdots & & \vdots \\ \sigma_{xny_1} & \sigma_{xny_2} & \cdots & \sigma_{xnym} \end{bmatrix} \qquad (1.2\text{-}12)$$

其中 $\sigma_{x_iy_j} = E\{(X_i-\mu_{x_i})(Y_j-\mu_{y_j})\}$。

协方差阵 $Cov(\boldsymbol{X},\boldsymbol{Y})$ 是随机矢量 $\boldsymbol{X}$ 的各分量对其均值的偏差，与随机矢量 $\boldsymbol{Y}$ 的各分量对其均值的偏差乘积平均值的矩阵。

（4） 条件数学期望

在 $X_1=x_1$，$X_2=x_2$，$\cdots$，$X_{n-1}=x_{n-1}$ 的条件下，X_n 的条件数学期望为

$$E\{X_n|x_1,x_2,\cdots,x_{n-1}\} \triangleq \int_{-\infty}^{\infty} x_n(x_n|x_1,x_2,\cdots,x_{n-1})dx_n \qquad (1.2\text{-}13)$$

同样，在随机矢量 $\boldsymbol{Y}=\boldsymbol{y}=(y_1y_2\cdots y_m)^T$ 条件下，随机矢量 $\boldsymbol{X}=(X_1X_2\cdots X_n)^T$ 的条件数学期望为

$$E\{\boldsymbol{X}|\boldsymbol{y}\} = \begin{bmatrix} E\{X_1|\boldsymbol{y}\} \\ E\{X_2|\boldsymbol{y}\} \\ \vdots \\ E\{X_n|\boldsymbol{y}\} \end{bmatrix} = \begin{bmatrix} E\{X_1|y_1,y_2,\cdots,y_m\} \\ E\{X_2|y_1,y_2,\cdots,y_m\} \\ \vdots \\ E\{X_n|y_1,y_2,\cdots,y_m\} \end{bmatrix} \qquad (1.2\text{-}14)$$

条件数学期望是滤波理论中的一个重要数据，它的统计意义留待下一章说明。

§1.3 随机变量线性变换的概率分布

滤波问题中所遇到的随机变量常常都是另一些随机变量的线性函数，因此计算一个或 n 个随机变量的线性变换的概率分布在滤波问题中是有实际意义的。

先考虑一个随机变量的线性变换的概率分布。

设随机变量 X 的分布密度是 $f_x(x)$，研究随机变量

$$Y = aX + b \quad (a,b \text{ 为非随机数}, \ a \neq 0)$$

的分布密度 $f_y(y)$。

由于 Y 的分布函数

$$F_y(y) = P(Y \leqslant y) = P(aX + b \leqslant y)$$

(i) 如果 $a>0$，则有

$$F_y(y) = P(aX + b \leqslant y) = P\left(X \leqslant \frac{y-b}{a}\right) = \int_{-\infty}^{\frac{y-b}{a}} f_x(x)dx$$

两边对 y 求导，得随机变量 y 的分布密度

$$f_y(y) = F'_y(y) = f_x\left(\frac{y-b}{a}\right)\frac{1}{a}$$

(ii) 如果 $a<0$，则有

$$F_y(y) = P(aX + b \leqslant y) = P\left(X \geqslant \frac{y-b}{a}\right) = \int_{\frac{y-b}{a}}^{\infty} f_x(x)dx$$

两边对 y 求导，得随机变量 Y 的分布密度

$$f_y(y) = F'_y(y) = -f_x\left(\frac{y-b}{a}\right)\frac{1}{a}$$

综合 (i)(ii) 两种情况下的结果，得随机变量 $Y = aX + b$ 的分布密度

$$f_y(y) = f_x\left(\frac{y-b}{a}\right)\frac{1}{|a|} \qquad (a \neq 0) \tag{1.3-1}$$

下面考虑随机矢量的线性变换的概率分布。

设 n 维随机矢量 $\boldsymbol{X}$ 的联合分布密度为 $f_x(\boldsymbol{x})$，研究 n 维随机矢量

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$$

的联合分布密度 $f_y(\boldsymbol{y})$，其中 $\boldsymbol{A}$ 为 $n \times n$ 非随机的可逆阵，$\boldsymbol{b}$ 为 n 维非随机矢量。

由逆变换 $\boldsymbol{X} = \boldsymbol{A}^{-1}(\boldsymbol{Y} - \boldsymbol{b})$，可以确定出一个 n 维空间域 $I(\boldsymbol{y}, \boldsymbol{A}, \boldsymbol{b})$，使得

$$-\infty < \boldsymbol{Y} \leqslant \boldsymbol{y} \Longleftrightarrow \boldsymbol{X} \in I(\boldsymbol{y}, \boldsymbol{A}, \boldsymbol{b})$$

于是 $\boldsymbol{Y}$ 的分布函数为

$$F_y(\boldsymbol{y}) = P(\boldsymbol{Y} \leqslant \boldsymbol{y}) = P\{\boldsymbol{X} \in I(\boldsymbol{y}, \boldsymbol{A}, \boldsymbol{b})\}$$

以分布密度的积分表示，上式可写成

$$\int_{-\infty}^{y} f_y(\boldsymbol{y})d\boldsymbol{y} = \int_{I(\boldsymbol{y}, \boldsymbol{A}, \boldsymbol{b})} f_x(\boldsymbol{x})d\boldsymbol{x}$$

对后一积分进行变量置换，令

$$\boldsymbol{x}=\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{b}),$$

$$d\boldsymbol{x}=\left|det\left[\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}}\right]\right|d\boldsymbol{y}=|det\boldsymbol{A}^{-1}|d\boldsymbol{y}=\frac{1}{|det\boldsymbol{A}|}d\boldsymbol{y}$$

积分式改变为

$$\int_{-\infty}^{y} f_y(\boldsymbol{y})d\boldsymbol{y}=\int_{-\infty}^{y} f_x[\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{b})]\frac{1}{|det\boldsymbol{A}|}d\boldsymbol{y}$$

两边对 $\boldsymbol{y}$ 求导，得随机矢量$\boldsymbol{Y}$的联合分布密度

$$f_y(\boldsymbol{y})=f_x[\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{b})]\frac{1}{|det\boldsymbol{A}|}\qquad (det\boldsymbol{A}\neq 0)\qquad (1.3\text{-}2)$$

比较式（1.3-1）与式（1.3-2），可以看出一维随机变量与多维随机变量线性变换的概率密度函数，形式完全相似。显然，式（1.3-2）只适用于$\boldsymbol{A}$为可逆的情况，如果$\boldsymbol{A}$不可逆，就不能这样确定 $f_y(\boldsymbol{y})$。

§1.4 正态随机变量与正态随机矢量

在滤波和其它的一些实践问题中，最常遇到的是这样的一类随机变量X，它的可能值分布在整个无穷区间（$-\infty,\infty$）上，概率密度函数 $f(x)$ 在某一点μ处具有最大值，越远离μ点概率密度就越小，它的几何图形是一条单峰且对称于$x=\mu$轴的对称曲线，如图 1.1-1 所示。其对应的函数解析式是

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\qquad (1.4\text{-}1)$$

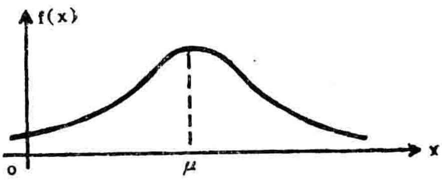

图 1.1-1 正态概率密度函数

具有这样的概率密度函数的随机变量X，称为正态（或高斯）随机变量，记作 $X\sim N(\mu,\sigma^2)$。

根据数学期望与方差的定义(1.1-6)及(1.1-7)，可以证明正态概率密度函数（1.4-1）中的参数

$$\mu=E\{X\},\quad \sigma^2=Var(X)\qquad (1.4\text{-}2)$$

又由式（1.3-1），正态随机变量X的线性函数$Y=aX+b$的概率密度函数为

$$f_y(y)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{1}{2\sigma^2}\left[\frac{y-b}{a}-\mu\right]^2}\frac{1}{|a|}$$

$$=\frac{1}{\sqrt{2\pi}\,|a|\sigma}e^{-\frac{1}{2a^2\sigma^2}[y-(a\mu+b)]^2}$$

由此可得出正态随机变量的一个重要性质：正态随机变量X 的线性函数$Y=aX+b$仍然是一个正态随机变量，其数学期望与方差分别是

$$E\{Y\}=a\mu+b,\quad Var(Y)=a^2\sigma^2\qquad (1.4\text{-}3)$$

同样，如果n维随机矢量$\boldsymbol{X}=(X_1X_2\cdots X_n)^T$的联合概率密度为

$$f(\boldsymbol{x})=\frac{1}{\sqrt{(2\pi)^n|\boldsymbol{V}_x|}}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_x)^T\boldsymbol{V}_x^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_x)} \tag{1.4-4}$$

其中 $\boldsymbol{V}_x$ 为 n 阶对称的正定阵①，则称 $\boldsymbol{X}=(X_1X_2\cdots X_n)^T$ 为 n 维正态随机矢量，记作 $\boldsymbol{X}\sim N(\boldsymbol{\mu}_x,\boldsymbol{V}_x)$。可以证明，式（1.4-4）中的

$$\boldsymbol{\mu}_x=E\{\boldsymbol{X}\}=\begin{bmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_n\end{bmatrix},\quad \boldsymbol{V}_x=Var(\boldsymbol{X})=\begin{bmatrix}\sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n}\\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n}\\ \vdots & \vdots & & \vdots\\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2\end{bmatrix}$$

$\boldsymbol{X}$的特征函数为

$$\begin{aligned}\Phi_x(\boldsymbol{S})&=E\{e^{i\boldsymbol{X}^T\boldsymbol{S}}\}=\int_{-\infty}^{\infty}e^{i\boldsymbol{x}^T\boldsymbol{S}}f(\boldsymbol{x})d\boldsymbol{x}\\ &=e^{\{i\boldsymbol{\mu}_x^T\boldsymbol{S}-\frac{1}{2}\boldsymbol{S}^T\boldsymbol{V}_x\boldsymbol{S}\}}\end{aligned} \tag{1.4-5}$$

由式（1.4-5）可以证明，正态随机矢量也具有一个重要的性质（见附录B）：正态随机矢量$\boldsymbol{X}$的线性变换 $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}+\boldsymbol{b}$ 仍然是一个正态随机矢量，其数学期望与方差分别是

$$E\{\boldsymbol{Y}\}=\boldsymbol{A}E\{\boldsymbol{X}\}+\boldsymbol{b},\quad Var(\boldsymbol{Y})=\boldsymbol{A}Var(\boldsymbol{X})\boldsymbol{A}^T \tag{1.4-6}$$

下面讨论正态随机矢量的条件数学期望的一个重要性质。

设有 $n+m$ 维正态随机矢量

$$\boldsymbol{Z}=\begin{bmatrix}\boldsymbol{X}\\ \boldsymbol{Y}\end{bmatrix}$$

其中$\boldsymbol{X}$为 n 维正态随机矢量，$\boldsymbol{Y}$为 m 维正态随机矢量，其数学期望矢量与方差阵分别是

$$E\{\boldsymbol{Z}\}=\begin{bmatrix}E\{\boldsymbol{X}\}\\ E\{\boldsymbol{Y}\}\end{bmatrix}=\begin{bmatrix}\boldsymbol{\mu}_x\\ \boldsymbol{\mu}_y\end{bmatrix}$$

$$\boldsymbol{V}_z=Var(\boldsymbol{Z})=\begin{bmatrix}Var(\boldsymbol{X}) & Cov(\boldsymbol{X},\boldsymbol{Y})\\ Cov(\boldsymbol{Y},\boldsymbol{X}) & Var(\boldsymbol{Y})\end{bmatrix}=\begin{bmatrix}\boldsymbol{V}_x & \boldsymbol{V}_{xy}\\ \boldsymbol{V}_{xy} & \boldsymbol{V}_y\end{bmatrix}$$

应用附录A分块矩阵的求逆法则，有

$$\boldsymbol{V}_z=\begin{bmatrix}\boldsymbol{I}_n & \boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\\ \boldsymbol{O} & \boldsymbol{I}_m\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx} & \boldsymbol{O}\\ \boldsymbol{V}_{yx} & \boldsymbol{V}_y\end{bmatrix}$$

$$|\boldsymbol{V}_z|=|\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx}|\,|\boldsymbol{V}_y|=|\boldsymbol{R}|\,|\boldsymbol{V}_y|$$

（其中$\boldsymbol{R}\triangleq\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx}$）

$$\begin{aligned}\boldsymbol{V}_z^{-1}&=\begin{bmatrix}(\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx})^{-1} & -(\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx})^{-1}\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\\ -\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx}(\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx})^{-1} & \boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx}(\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx})^{-1}\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}+\boldsymbol{V}_y^{-1}\end{bmatrix}\\ &=\begin{bmatrix}\boldsymbol{R}^{-1} & -\boldsymbol{R}^{-1}\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\\ -\boldsymbol{V}_y^{-1}\boldsymbol{V}_{xy}^T\boldsymbol{R}^{-1} & \boldsymbol{V}_y^{-1}\boldsymbol{V}_{xy}^T\boldsymbol{R}^{-1}\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}+\boldsymbol{V}_y^{-1}\end{bmatrix}\\ &=\begin{bmatrix}\boldsymbol{I}_n\\ -\boldsymbol{V}_y^{-1}\boldsymbol{V}_{xy}^T\end{bmatrix}\boldsymbol{R}^{-1}[\boldsymbol{I}_n\ \ -\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}]+\begin{bmatrix}\boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{V}_y^{-1}\end{bmatrix}\end{aligned}$$

其中 $\boldsymbol{R}=\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{yx}$，于是根据正态随机矢量联合分布密度(概率密度)函数（1.4-4)，有

---

① 当$\boldsymbol{V}_x$只是非负定阵时，则需要改用特征函数来定义 n 维正态随机矢量(见参考文献[26]第一册)。

$$f(\boldsymbol{Z})=f(\boldsymbol{x},\boldsymbol{y})=\frac{1}{\sqrt{(2\pi)^{n+m}|\boldsymbol{V}_z|}}e^{-\frac{1}{2}[(\boldsymbol{x}-\boldsymbol{\mu}_x)^T(\boldsymbol{y}-\boldsymbol{\mu}_y)^T)\boldsymbol{V}_z^{-1}\begin{bmatrix}\boldsymbol{x}-\boldsymbol{\mu}_x\\ \boldsymbol{y}-\boldsymbol{\mu}_y\end{bmatrix}}$$

$$=\frac{1}{\sqrt{(2\pi)^{n+m}|\boldsymbol{R}||\boldsymbol{V}_y|}}e^{-\frac{1}{2}[\boldsymbol{x}-\boldsymbol{\mu}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_y)]^T\boldsymbol{R}^{-1}[\boldsymbol{x}-\boldsymbol{\mu}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_y)]}$$

$$\cdot\ e^{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}_y)^T\boldsymbol{V}_y^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_y)}$$

$$f_y(\boldsymbol{y})=\frac{1}{\sqrt{(2\pi)^m|\boldsymbol{V}_y|}}e^{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}_y)^T\boldsymbol{V}_y^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_y)}$$

再由分布密度函数的贝叶斯公式，即可得出在 $\boldsymbol{Y}=\boldsymbol{y}$ 条件下$\boldsymbol{X}$的条件分布密度

$$f(\boldsymbol{x}|\boldsymbol{y})=\frac{f(\boldsymbol{x},\boldsymbol{y})}{f_y(\boldsymbol{y})}$$

$$=\frac{1}{\sqrt{(2\pi)^n|\boldsymbol{R}|}}e^{-\frac{1}{2}[\boldsymbol{x}-\boldsymbol{\mu}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_y)]^T\boldsymbol{R}^{-1}[\boldsymbol{x}-\boldsymbol{\mu}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_y)]} \tag{1.4-7}$$

由上式可以看出，在 $\boldsymbol{Y}=\boldsymbol{y}$ 条件下$\boldsymbol{X}$仍然服从于正态分布，其条件数学期望矢量与条件方差阵分别为

$$\begin{aligned}E\{\boldsymbol{X}|\boldsymbol{y}\}&=\boldsymbol{\mu}_x+\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_y)\\&=E\{\boldsymbol{X}\}+Cov(\boldsymbol{X},\boldsymbol{Y})[Var(\boldsymbol{Y})]^{-1}(\boldsymbol{y}-E\{\boldsymbol{Y}\})\end{aligned} \tag{1.4-8}$$

$$\begin{aligned}Var(\boldsymbol{X}|\boldsymbol{y})&=\boldsymbol{R}=\boldsymbol{V}_x-\boldsymbol{V}_{xy}\boldsymbol{V}_y^{-1}\boldsymbol{V}_{xy}^T\\&=Var(\boldsymbol{X})-Cov(\boldsymbol{X},\boldsymbol{Y})[Var(\boldsymbol{Y})]^{-1}Cov(\boldsymbol{Y},\boldsymbol{X})\end{aligned} \tag{1.4-9}$$

由式(1.4-8)得到一个关于正态随机矢量的条件数学期望的重要结论：当 $(\boldsymbol{X}^T\boldsymbol{Y}^T)^T$ 为正态随机矢量时，在 $\boldsymbol{Y}=\boldsymbol{y}$ 条件下，$\boldsymbol{X}$的条件数学期望 $E\{\boldsymbol{X}|\boldsymbol{y}\}$ 为 $\boldsymbol{y}$ 的线性 函数。以后将会看到，这个结论在滤波理论中具有十分重要的作用。

§1.5 正 交 投 影

1.5.1 不相交、正交与独立随机矢量

设 $\boldsymbol{X},\boldsymbol{Y}$ 是两个随机矢量，如果

$$\begin{aligned}Cov(\boldsymbol{X},\boldsymbol{Y})&=E\{(\boldsymbol{X}-\boldsymbol{\mu}_x)(\boldsymbol{Y}-\boldsymbol{\mu}_y)^T\}\\&=E\{\boldsymbol{X}\boldsymbol{Y}^T\}-\boldsymbol{\mu}_x\boldsymbol{\mu}_y^T=\boldsymbol{O}\end{aligned} \tag{1.5-1}$$

或

$$E\{\boldsymbol{X}\boldsymbol{Y}^T\}=E\{\boldsymbol{X}\}E^T\{\boldsymbol{Y}\} \tag{1.5-1'}$$

则称 $\boldsymbol{X},\boldsymbol{Y}$ 为不相关的随机矢量；如果

$$E\{\boldsymbol{X}\boldsymbol{Y}^T\}=\boldsymbol{O} \tag{1.5-2}$$

则称 $\boldsymbol{X},\boldsymbol{Y}$ 为正交的随机矢量；如果 $(\boldsymbol{X},\boldsymbol{Y})$ 的联合分布密度

$$f(\boldsymbol{x},\boldsymbol{y})=f_1(\boldsymbol{x})f_2(\boldsymbol{y}) \tag{1.5-3}$$

其中 $f_1(\boldsymbol{x})$ 为 $\boldsymbol{X}$ 的分布密度，$f_2(\boldsymbol{y})$ 为 $\boldsymbol{Y}$ 的分布密度，则称 $\boldsymbol{X},\boldsymbol{Y}$ 为相互独立的随机矢量。

从式(1.5-1)、(1.5-2)及(1.5-3)可以看出，这三个重要概念之间具有以下关系：

(1) 两个相互独立的随机矢量，一定是不相关的(但逆命题不成立)。

事实上，如果

$$f(\boldsymbol{x},\boldsymbol{y})=f_1(\boldsymbol{x})f_2(\boldsymbol{y})$$

则有

$$\begin{aligned}E\{\boldsymbol{X}\boldsymbol{Y}^T\}&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\boldsymbol{x}\boldsymbol{y}^Tf(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y}\\&=\int_{-\infty}^{\infty}\boldsymbol{x}f_1(\boldsymbol{x})d\boldsymbol{x}\cdot\int_{-\infty}^{\infty}\boldsymbol{y}^Tf_2(\boldsymbol{y})d\boldsymbol{y}=E\{\boldsymbol{X}\}E^T\{\boldsymbol{Y}\}\end{aligned}$$

故 $\boldsymbol{X},\boldsymbol{Y}$ 又是不相关的。

(2) 两个正态不相关的随机矢量，一定是相互独立的。

事实上，如果 $\boldsymbol{X},\boldsymbol{Y}$ 分别是 n 维及 m 维随机矢量，且互不相关，根据式(1.5-1)，有

$$\boldsymbol{V}_{xy}=E\{(\boldsymbol{X}-\boldsymbol{\mu}_x)(\boldsymbol{Y}-\boldsymbol{\mu}_y)^T\}=E\{\boldsymbol{X}\boldsymbol{Y}^T\}-E\{\boldsymbol{X}\}E^T\{\boldsymbol{Y}\}=\boldsymbol{O}$$

又由于 $(\boldsymbol{X}^T\boldsymbol{Y}^T)^T$ 服从于正态分布，根据式(1.4-4)，$(\boldsymbol{X}^T\boldsymbol{Y}^T)^T$ 的联合分布密度为

$$f(\boldsymbol{x},\boldsymbol{y})=\frac{1}{\sqrt{(2\pi)^{n+m}|\boldsymbol{V}|}}e^{-\frac{1}{2}[(\boldsymbol{x}-\boldsymbol{\mu}_x)^T(\boldsymbol{y}-\boldsymbol{\mu}_y)^T]\boldsymbol{V}^{-1}\begin{bmatrix}\boldsymbol{x}-\boldsymbol{\mu}_x\\\boldsymbol{y}-\boldsymbol{\mu}_y\end{bmatrix}}$$

其中

$$\boldsymbol{V}=\begin{bmatrix}\boldsymbol{V}_x&\boldsymbol{V}_{xy}\\\boldsymbol{V}_{yx}&\boldsymbol{V}_y\end{bmatrix}=\begin{bmatrix}\boldsymbol{V}_x&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{V}_y\end{bmatrix},\quad\boldsymbol{V}^{-1}=\begin{bmatrix}\boldsymbol{V}_x^{-1}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{V}_y^{-1}\end{bmatrix}$$

于是有

$$\begin{aligned}f(\boldsymbol{x},\boldsymbol{y})&=\frac{1}{(\sqrt{2\pi})^{n+m}|\boldsymbol{V}_x||\boldsymbol{V}_y|}e^{-\frac{1}{2}[(\boldsymbol{x}-\boldsymbol{\mu}_x)^T(\boldsymbol{y}-\boldsymbol{\mu}_y)^T]\begin{bmatrix}\boldsymbol{V}_x^{-1}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{V}_y^{-1}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}-\boldsymbol{\mu}_x\\\boldsymbol{y}-\boldsymbol{\mu}_y\end{bmatrix}}\\&=\frac{1}{(\sqrt{2\pi})^n|\boldsymbol{V}_x|}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_x)^T\boldsymbol{V}_x^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_x)}\\&\quad\cdot\frac{1}{(\sqrt{2\pi})^m|\boldsymbol{V}_y|}e^{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}_y)^T\boldsymbol{V}_y^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_y)}\\&=f_1(\boldsymbol{x})f_2(\boldsymbol{y})\end{aligned}$$

故 $\boldsymbol{X},\boldsymbol{Y}$ 相互独立。

(3) 从式(1.5-1)及式(1.5-2)容易看出，对于两个数学期望为零的随机矢量，“不相关”与“正交”两个概念是一致的。对于两个数学期望为零的联合正态随机矢量，“不相关”、“正交”与“独立”三个概念都是一致的。

1.5.2 正交投影的定义[10]

定义 设 $\boldsymbol{X}$ 与 $\boldsymbol{Z}$ 分别为具有前二阶矩的 n 维和 m 维随机矢量，如果存在一个与 $\boldsymbol{X}$ 同维