

▶▶○○○▶○○○▶▶▶▶

离散信号的滤波

王欣 王德隽 著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

国家自然科学基金资助项目

离散信号的滤波

王欣 王德隽 著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

滤波是现代数字信号处理的重要研究内容,在信号分析、图像处理、模式识别、自动控制等领域得到了广泛应用。最佳滤波是指在给定某种准则的条件下,得到信号的最佳估计。本书首先在最小均方误差条件下,介绍 Wiener 滤波、Kalman 滤波和自适应滤波等一些线性滤波方法。然后讨论近年来受到人们日益重视的一些非线性滤波方法,即小波滤波、同态滤波、中值滤波、形态滤波等,它们都是现代数字信号处理领域的前沿研究课题,不但有重要的理论意义,而且有广阔的应用前景。

本书适合通信、信号处理、自动控制、人工智能等专业的师生与科研人员阅读。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

离散信号的滤波/王欣等著. —北京:电子工业出版社, 2002.6

ISBN 7-5053-7594-6

I. 离… II. 王… III. 离散信号—滤波技术 IV. TN713

中国版本图书馆 CIP 数据核字(2002)第 029616 号

责任编辑:竺南直

印 刷:北京四季青印刷厂

出版发行:电子工业出版社 <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销:各地新华书店

开 本:787×1092 1/16 印张:12 字数:307 千字

版 次:2002 年 6 月第 1 版 2002 年 6 月第 1 次印刷

印 数:4 000 册 定价:20.00 元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。联系电话:(010)68279077

前 言

滤波是现代数字信号处理的重要研究内容,它在信号分析、图像处理、模式识别、自动控制等领域得到了广泛应用。本书主要介绍用于离散信号的一些典型的滤波方法。

滤波一般要求在给定某种准则的条件下,得到信号的最佳估计。Wiener 滤波是最早提出的一种滤波方法,当信号混有白噪声时,可以在最小均方误差条件下得到信号的最佳估计。但是,由于求解 Wiener-Hoff 方程的复杂性,使得 Wiener 滤波实际应用起来很困难。不过 Wiener 滤波在理论上的意义是非常重要的,利用 Wiener 滤波的纯一步预测,可以求解信号的模型参数,进而获得著名的 Levinson 算法。

Kalman 滤波是 20 世纪 60 年代初提出的一种滤波方法。与 Wiener 滤波相似,它同样可以在最小均方误差条件下给出信号的最佳估计。所不同的是,这种滤波在时域中采用递推方式进行,因此速度快,便于实时处理,从而得到了广泛的应用。Kalman 滤波推广到二维,可以用于图像的去噪。

当假设 Wiener 滤波器的单位脉冲响应为有限长时,可以采用被称为自适应滤波的方法来得到滤波器的最佳响应。由于它避开了求解 Wiener-Hoff 方程,故给某些问题的解决带来了极大的方便。而由 Windrow 提出的这种自适应滤波方法的基本思想,对信号处理中一些方法的改进产生了积极深远的影响。

小波变换是一种用于信号分析的新方法,它的多分辨率特性使其在许多领域得到了应用。小波滤波就是利用信号和噪声在各尺度下的小波变换系数有所不同的特点,来对它们进行分离,从而达到去除噪声的目的。尽管已有一些小波滤波的方法被提出,目前它仍是人们非常关注的一个研究领域。

同态滤波主要用于解决信号和噪声之间不是相加、而是相乘关系时的滤波问题。另外,当信号和噪声之间为卷积关系时,在一定的条件下也可以利用同态滤波把信号有效地分离出来。由同态滤波理论引申出来的复时谱,已成为现代信号处理中极为重要的概念。

Wiener 滤波、Kalman 滤波和自适应滤波都是线性滤波。线性滤波的最大缺点就是在滤除噪声的同时,会造成信号边缘的模糊。中值滤波是在 20 世纪 70 年代提出的一种非线性滤波方法,它可以在最小绝对误差条件下,给出信号的最佳估计。这种滤波方法的优点,就是能够保持信号的边缘不模糊。另外,它对脉冲噪声也有良好的清除作用。但是,由于理论分析上的复杂性,这种方法至今还有一些问题需要解决。

形态滤波是建立在集合运算之上的一种非线性滤波方法。它除了用于滤除信号中的噪声外,还在图像分析中发挥了重要作用。近年来,其理论和应用的研究受到了人们的极大重视。

本书对上述各种滤波方法的基本原理进行了详细的介绍,力求做到理论完整、论述简洁、推导严密、分析正确。其中包括了著者多年来的一些研究成果。鉴于作者的学术水平有限,不足之处恳请读者指正。

该书的出版得到了国家自然科学基金项目的资助,电子工业出版社的竺南直编辑也提出了宝贵的建议,在这里我们表示衷心的感谢!

著 者
2001 年 10 月

目 录

第 1 章 离散随机信号	(1)
1.1 随机变量及统计特性	(1)
1.2 正交投影原理	(7)
1.3 离散随机信号	(10)
1.4 随机序列经过线性滤波器	(17)
1.5 最小相位滤波器	(19)
1.6 信号模型	(21)
1.7 随机变量的参数估计	(24)
第 2 章 Wiener 滤波	(28)
2.1 信号的滤波	(28)
2.2 Wiener-Hoff 方程	(30)
2.3 平稳序列的 Wiener 滤波	(33)
2.4 平稳序列的 Wiener 预测	(37)
2.5 Levinson-Durbin 算法	(42)
2.6 格型滤波器	(46)
2.7 Burg 算法	(52)
2.8 功率谱估计	(53)
第 3 章 Kalman 滤波	(59)
3.1 状态与观测方程	(59)
3.2 Kalman 滤波	(60)
3.3 有色噪声模型的 Kalman 滤波	(68)
第 4 章 自适应滤波	(73)
4.1 FIR Wiener 滤波器	(73)
4.2 LMS 自适应算法	(74)
4.3 步长因子 μ 的取值范围	(78)
4.4 学习曲线	(79)
4.5 失调	(81)
4.6 RLS 自适应滤波	(84)
4.7 自适应噪声抵消	(87)
4.8 自适应信道均衡器	(91)
第 5 章 小波滤波	(93)
5.1 连续小波变换	(93)
5.2 二进小波变换	(97)

5.3	离散二进小波变换	(99)
5.4	Daubechies 小波	(110)
5.5	双正交小波变换	(113)
5.6	信号的最大模重建	(119)
5.7	利用最大模重建滤除噪声	(121)
5.8	软门限去噪	(125)
第 6 章	同态滤波	(128)
6.1	同态滤波的定义	(128)
6.2	解相乘同态系统	(129)
6.3	解卷积同态系统	(131)
6.4	复时谱的计算	(133)
6.5	指数序列的复时谱	(136)
第 7 章	中值滤波	(143)
7.1	中值滤波的定义	(143)
7.2	中值滤波的门限分解算法	(146)
7.3	中值滤波的输出统计特性	(148)
7.4	多级中值滤波	(152)
7.5	序统计滤波	(156)
7.6	近均值滤波	(159)
7.7	Lee 滤波	(161)
7.8	梯度倒数加权滤波	(165)
第 8 章	形态滤波	(168)
8.1	形态学的基本运算和性质	(168)
8.2	Matheron 表示定理	(175)
8.3	一维信号的形态滤波	(177)
8.4	一维形态滤波的输出统计特性	(179)
参考文献		(185)

第 1 章 离散随机信号

在离散信号的滤波中，我们需要处理的对象一般都是随机信号经过采样量化后所得到的数字信号。因此，离散信号滤波方法的研究和随机信号的理论有着密切的联系。本章我们首先讨论了随机变量的统计特性，通过特征函数求解随机矢量高阶混合矩的方法及最佳投影原理；其次给出了离散随机信号自相关函数和功率谱密度的定义，以及它们经过线性系统后的变化，其中涉及到随机信号模型、最小相位滤波器等；最后对参数估计的基本方法进行了介绍。这些内容是研究离散信号滤波方法^[1]的基础。

1.1 随机变量及统计特性

令 x 为一随机变量，其概率密度函数为 $p(x)$ ，那么该随机变量的均值、方差及二阶矩分别由下面各式定义

$$\mu = E[x] = \int_{-\infty}^{\infty} xp(x)dx$$

$$\sigma^2 = E[(x - \mu_x)^2] = \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x)dx$$

$$E[x^2] = \int_{-\infty}^{\infty} x^2 p(x)dx$$

其中 $E[\]$ 表示取数学期望。容易证明

$$\sigma^2 = E[x^2] - \mu^2$$

上述统计量的重要性表现在它们是进行最佳滤波器设计时所需要知道的基本条件，并且这些统计量可以通过实际的测量数据计算出来。

具有 Gauss 分布的随机变量 x 的概率密度函数 $p(x)$ ，由下式给出

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

其中 μ 和 σ^2 分别为随机变量 x 的均值和方差。由此可以看出，Gauss 分布随机变量的概率密度函数完全由其均值和方差这两个统计量所决定。

Gauss 分布在信号的滤波理论中是非常重要的，其原因是我们一般假设有用信号中所混有的噪声具有 Gauss 分布。考虑到噪声往往是由许多独立的噪声源经叠加后所形成的，那么根据大数定理，这一假设是合理的。

对于随机变量的均值和方差，下面的 Chebyshev 不等式给出了它们之间的重要联系。

定理 1.1.1 若随机变量 x 具有有限方差 σ_x^2 , 则对于任一正数 ε 有

$$p[|x - \mu_x| \geq \varepsilon] \leq \frac{\sigma_x^2}{\varepsilon^2} \quad (1.1.1)$$

证明: 设 $p(x)$ 是随机变量 x 的概率密度函数, 则

$$\begin{aligned} p[|x - \mu_x| \geq \varepsilon] &= \int_{|x - \mu_x| \geq \varepsilon} p(x) dx \leq \int_{|x - \mu_x| \geq \varepsilon} \frac{(x - \mu_x)^2}{\varepsilon^2} p(x) dx \\ &\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x) dx = \frac{\sigma_x^2}{\varepsilon^2} \end{aligned}$$

定理证毕。

Chebyshev 不等式在概率论中有着重要作用。它表明当方差愈小时, 事件 $|x - \mu_x| \geq \varepsilon$ 的概率愈小, 即随机变量的取值密集于其均值的周围。因而方差可用来作为描述随机变量对于它的均值的分散指标。

应用对立事件的概念, Chebyshev 不等式又可以写成

$$p[|x - \mu_x| < \varepsilon] \geq 1 - \frac{\sigma_x^2}{\varepsilon^2} \quad (1.1.2)$$

该不等式给出了在随机变量的分布函数未知, 而方差已知的情况下, 对事件 $|x - \mu_x| < \varepsilon$ 的概率的一种估计方法。

当把随机变量看成一维的随机矢量时, 那么由两个不同的随机变量组成的随机变量对 $\mathbf{x} = (x_1, x_2)$ 形成一个二维随机矢量。相对于随机变量而言, 二维随机矢量的统计描述更为复杂, 即要求知道它们的联合概率密度函数 $p(x_1, x_2)$ 。所给定的两个随机变量之间可以是非独立的, 也可以是独立的。两个随机变量之间是非独立的是指当随机变量 x_2 的值给定后, 它会直接影响 x_1 的取值。对于此, 我们用条件概率密度函数 $p(x_1 | x_2)$ 来进行描述, 并有下列的 **Bayes 公式**

$$p(x_1, x_2) = p(x_1 | x_2) p(x_2) = p(x_2 | x_1) p(x_1)$$

当两个随机变量之间为独立时, 则 $p(x_1 | x_2) = p(x_1)$ 或 $p(x_2 | x_1) = p(x_2)$ 。由上式可知此时

$$p(x_1, x_2) = p(x_1) p(x_2)$$

反之, 当该式成立时, 两个随机变量之间也是独立的。

随机变量 x_1 和 x_2 之间的关联特性也可以用 **互相关** 来进行描述, 它的定义由下式给出

$$E[x_1 x_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2) dx_1 dx_2 \quad (1.1.3)$$

当随机变量 x_1 和 x_2 之间为相互独立时, $E[x_1 x_2] = E[x_1] E[x_2]$ 。

我们可以用列向量

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

表示 N 维的随机矢量，其中 $x_1, x_2, \dots, x_N$ 为 N 个不同的随机变量。为了对该 N 维随机矢量进行描述，需要知道 N 维联合概率密度函数 $p(\mathbf{x}) = p(x_1, x_2, \dots, x_N)$ 。随机矢量的均值、相关阵和协方差阵，分别由下面各式定义

$$\mu_x = E[\mathbf{x}]$$

$$\Phi_{xx} = E[\mathbf{x}\mathbf{x}^\tau]$$

$$\mathbf{C}_{xx} = E[(\mathbf{x} - \mu_x)(\mathbf{x} - \mu_x)^\tau]$$

这里，上标“ τ ”表示对矩阵的转置。

显然，相关阵和协方差阵是对称矩阵，如果随机矢量的均值为零则它们相等。下面，我们要证明相关阵的非负定特性。

定理 1.1.2 N 维随机矢量 $\mathbf{x}$ 的自相关矩阵 Φ_{xx} 是非负定的。

证明：令 $\mathbf{h} = [h(0)h(1)\cdots h(N-1)]^\tau$ 为由任意 N 个实数组成的列向量。这样，为了证明 Φ_{xx} 是非负定的，只要证明 $\mathbf{h}^\tau \Phi_{xx} \mathbf{h} \geq 0$ 便可。实际上根据定义可知

$$\begin{aligned} \mathbf{h}^\tau \Phi_{xx} \mathbf{h} &= \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} h(n)h(k)E[x(n)x(k)] \\ &= E\left[\sum_{n=0}^{N-1} \sum_{k=0}^{N-1} h(n)h(k)x(n)x(k)\right] \\ &= E\left\{\left[\sum_{n=0}^{N-1} h(n)x(n)\right]^2\right\} \geq 0 \end{aligned}$$

故定理得证。

与一维情形相似， N 维 Gauss 随机矢量的概率密度函数也是由它的均值和协方差矩阵所决定，并可写成

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} (\det \mathbf{C}_{xx})^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mu_x)^\tau \mathbf{C}_x^{-1}(\mathbf{x} - \mu_x)\right] \quad (1.1.4)$$

Gauss 随机矢量有下述重要特性。

性质 1.1.1 一 Gauss 随机矢量经过线性变换后，仍然是 Gauss 随机矢量。

证明：令 $\mathbf{x}$ 是均值和协方差阵分别为 μ_x 和 $\mathbf{C}_{xx}$ 的随机矢量。经过线性变换 $\xi = \mathbf{L}\mathbf{x}$ 变换后，这里 $\mathbf{L}$ 为一个非奇异矩阵，其均值和协方差阵分别为 $\mu_\xi = \mathbf{L}\mu_x$ 和 $\mathbf{C}_{\xi\xi} = \mathbf{L}\mathbf{C}_x\mathbf{L}^\tau$ 。设变换前后的随机矢量 $\mathbf{x}$ 和 ξ 的概率密度函数分别为 $p_x(\mathbf{x})$ 和 $p_\xi(\xi)$ ，那么有 $p_x(\mathbf{x})d\mathbf{x}_1 \cdots d\mathbf{x}_N = p_\xi(\xi)d\xi_1 \cdots d\xi_N$ 。利用线性变换的 Jacobi 关系式

$d\xi_1 \cdots d\xi_N = |\det L| dx_1 \cdots dx_N$, 便可以得到 $p_\xi(\xi) = p_x(x)/|\det L|$ 。注意到

$$\begin{aligned}(\xi - \mu_\xi)^T C_{\xi\xi}^{-1} (\xi - \mu_\xi) &= (x - \mu_x)^T L^T (LC_{xx}L^T)^{-1} L(x - \mu_x) \\&= (x - \mu_x)^T C_x^{-1} (x - \mu_x)\end{aligned}$$

及

$$\det C_{\xi\xi} = \det(LC_{xx}L^T) = (\det L)^2 \det C_{xx}$$

最后由式 (1.1.4) 可得

$$p(\xi) = \frac{1}{(2\pi)^{N/2} (\det C_\xi)^{1/2}} \exp\left[-\frac{1}{2}(\xi - \mu_\xi)^T C_{\xi\xi}^{-1} (\xi - \mu_\xi)\right]$$

这样, 性质 1.1.1 便得到了证明。

当两个均值为零的随机矢量不相关时, 它们并不一定相互独立。但对 Gauss 随机矢量, 其独立性是肯定的, 下面就对此加以证明。

性质 1.1.2 两个维数分别为 M, N , 均值为零的随机矢量 x, y , 如果它们不相关且是联合 Gauss 分布的, 则它们之间相互独立。

证明: 令 $z = \begin{bmatrix} x \\ y \end{bmatrix}$, 则由已知条件, 随机矢量 z 的概率密度函数可写为

$$p(z) = \frac{1}{(2\pi)^{(M+N)/2} (\det \Phi_z)^{1/2}} \exp\left[-\frac{1}{2}z^T \Phi_z^{-1} z\right]$$

这里, 自相关矩阵

$$\Phi_z = E[zz^T] = \begin{bmatrix} E[xx^T] & E[xy^T] \\ E[yx^T] & E[yy^T] \end{bmatrix} = \begin{bmatrix} \Phi_{xx} & \Phi_{xy} \\ \Phi_{yx} & \Phi_{yy} \end{bmatrix}$$

由随机矢量 x 和 y 不相关的条件, 可得 $\Phi_{xy} = E[xy^T] = 0$ 。因此

$$z^T \Phi_z^{-1} z = E[zz^T] = \begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} \Phi_{xx}^{-1} & 0 \\ 0 & \Phi_{yy}^{-1} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^T \Phi_{xx}^{-1} x + y^T \Phi_{yy}^{-1} y$$

另外, 根据 $\Phi_{xy} = 0$ 还可以推出 $\det \Phi_z = (\det \Phi_{xx})(\det \Phi_{yy})$ 。把它们代入随机矢量 z 的概率密度函数表示式中, 就可以得到 $p(z) = p(x)p(y)$ 。这样我们就证明了随机矢量 x 和 y 之间是独立的。

在求随机变量的各阶矩, 特别是高阶矩时, 特征函数是一简便有效的数学工具。当给定随机变量 x 及其概率密度函数 $p(x)$ 后, 则称

$$\phi(v) = E[e^{jvx}] = \int_{-\infty}^{\infty} p(x)e^{jvx} dx \quad (1.1.5)$$

为随机变量 x 的**特征函数**, 其中 v 为实参变量。从上式知道,

$$|\phi(v)| = \left| \int_{-\infty}^{\infty} p(x) e^{jvx} dx \right| \leq \int_{-\infty}^{\infty} p(x) dx = 1$$

由于式 (1.1.5) 可以看成为函数 $p(-x)$ 的 Fourier 变换, 故由逆 Fourier 变换公式可得

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(v) e^{-jvx} dv \quad (1.1.6)$$

下面我们就证明若随机变量 x 有 n 阶绝对矩, 则

$$E[x^k] = j^{-k} \phi^{(k)}(0) \quad (1.1.7)$$

其中, $k < n$ 。

实际上由式 (1.1.5) 可知

$$\phi^{(k)}(v) = \int_{-\infty}^{\infty} j^k x^k p(x) e^{jvx} dx$$

当随机变量有 n 阶绝对矩时, 上式积分收敛, 故

$$j^{-k} \phi^{(k)}(0) = \int_{-\infty}^{\infty} x^k p(x) dx$$

这样, 式 (1.1.7) 便得到了证明。

同样, 对于 N 维随机矢量 $\mathbf{x}$, 其特征函数可类似定义为

$$\begin{aligned} \phi(v_1, \dots, v_N) &= E[\exp(j \sum_{k=1}^N v_k x_k)] \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_1, \dots, x_N) \exp[j(v_1 x_1 + \dots + v_N x_N)] dx_1 \dots dx_N \end{aligned} \quad (1.1.8)$$

其中 $p(x_1, \dots, x_N)$ 为该随机矢量的联合概率密度函数。如果令 $\mathbf{v} = (v_1, \dots, v_N)^T$, $\mathbf{x} = (x_1, \dots, x_N)^T$ 则上式也可写成矢量形式

$$\phi(\mathbf{v}) = \int_{R^N} p(\mathbf{x}) \exp(j\mathbf{v}^T \mathbf{x}) d\mathbf{x}$$

而与式 (1.1.6) 相对应的逆变换公式则可写为

$$p(\mathbf{x}) = \frac{1}{(2\pi)^N} \int_{R^N} \phi(\mathbf{v}) \exp(-j\mathbf{v}^T \mathbf{x}) d\mathbf{v} \quad (1.1.9)$$

对于 Gauss 随机矢量, 有下面的定理。

定理 1.1.3 由式 (1.1.4) 所给出的 N 维 Gauss 分布的特征函数为

$$\phi(\mathbf{v}) = \exp(j\mu_x^T \mathbf{v} - \frac{1}{2} \mathbf{v}^T \mathbf{C}_{xx} \mathbf{v}) \quad (1.1.10)$$

证明: 首先根据定义

$$\begin{aligned}\phi(\mathbf{v}) &= \int_{R^N} p(\mathbf{x}) \exp(j\mathbf{v}^T \mathbf{x}) d\mathbf{x} \\ &= \frac{1}{(2\pi)^{N/2} (\det \mathbf{C}_{xx})^{1/2}} \int_{R^N} \exp(j\mathbf{v}^T \mathbf{x}) \exp[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_x)^T \mathbf{C}_{xx}^{-1}(\mathbf{x} - \boldsymbol{\mu}_x)] d\mathbf{x}\end{aligned}$$

设存在非奇异矩阵 $\mathbf{L}$, 使 $\mathbf{C}_{xx} = \mathbf{L}\mathbf{L}^T$ 。如果令 $\mathbf{y} = \mathbf{L}^{-1}(\mathbf{x} - \boldsymbol{\mu}_x)$, $\mathbf{s} = \mathbf{L}^T \mathbf{v}$, 有

$$\begin{aligned}j\mathbf{v}^T \mathbf{x} - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_x)^T \mathbf{C}_{xx}^{-1}(\mathbf{x} - \boldsymbol{\mu}_x) &= j\mathbf{v}^T \boldsymbol{\mu}_x + j\mathbf{v}^T \mathbf{L}\mathbf{y} - \frac{1}{2}\mathbf{y}^T \mathbf{y} \\ &= j\mathbf{v}^T \boldsymbol{\mu}_x + j\mathbf{s}^T \mathbf{y} - \frac{1}{2}\mathbf{y}^T \mathbf{y} \\ &= j\mathbf{v}^T \boldsymbol{\mu}_x - \frac{1}{2} \sum_{k=1}^N (y_k - js_k)^2 - \frac{1}{2}\mathbf{s}^T \mathbf{s} \\ &= j\mathbf{v}^T \boldsymbol{\mu}_x - \frac{1}{2} \sum_{k=1}^N (y_k - js_k)^2 - \frac{1}{2}\mathbf{v}^T \mathbf{C}_{xx} \mathbf{v}\end{aligned}$$

代入式 (1.1.10) 后, 并注意 $|\det \mathbf{L}| = |\det \mathbf{C}_{xx}|^{1/2}$ 得

$$\begin{aligned}\phi(\mathbf{v}) &= \frac{\exp(j\mathbf{v}^T \boldsymbol{\mu}_x - \frac{1}{2}\mathbf{v}^T \mathbf{C}_{xx} \mathbf{v})}{(2\pi)^{N/2} (\det \mathbf{C}_{xx})^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp[-\frac{1}{2} \sum_{k=1}^N (y_k - js_k)^2] |\det \mathbf{C}_{xx}|^{1/2} dy_1 \cdots dy_N \\ &= \exp(j\mathbf{v}^T \boldsymbol{\mu}_x - \frac{1}{2}\mathbf{v}^T \mathbf{C}_{xx} \mathbf{v}) \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-\frac{u^2}{2}) du \right]^N \\ &= \exp(j\mathbf{v}^T \boldsymbol{\mu}_x - \frac{1}{2}\mathbf{v}^T \mathbf{C}_{xx} \mathbf{v})\end{aligned}$$

这样, 定理得到了证明。

设 $\{x_1, \dots, x_N\}$ 为零均值的 Gauss 随机矢量, 利用特征函数的定义, 容易证明它的高阶混合中心矩

$$E[x_1^{b_1} \cdots x_N^{b_N}] = (-j)^B \frac{\partial^B}{\partial v_1^{b_1} \cdots \partial v_N^{b_N}} \phi(v_1 \cdots v_N) \Big|_{v_1 = \cdots = v_N = 0} \quad (1.1.11)$$

这里, $B = b_1 + \cdots + b_N$ 。再考虑到零均值的 Gauss 随机矢量 $\mathbf{x}$ 的特征函数为

$$\begin{aligned}\phi(\mathbf{v}) &= \exp(-\frac{1}{2}\mathbf{v}^T \mathbf{C}_{xx} \mathbf{v}) \\ &= \exp(-\frac{1}{2} \sum_{i=1}^N \sum_{k=1}^N v_k c_{ki} v_i)\end{aligned}$$

式中的 c_{ki} 为协方差阵 $\mathbf{C}_{xx}$ 的第 ki 个元素。利用指数展开式, 得

$$\phi(\mathbf{v}) = \sum_{r=0}^{\infty} \frac{(-1)^r}{2^r r!} \left[\sum_{i=1}^N \sum_{k=1}^N v_k c_{ki} v_i \right]^r \quad (1.1.12)$$

由上式能够看出, $\phi(\mathbf{v})$ 可以展开成一个级数, 其中每一项由各个 $v_i, i=1, \dots, N$ 的方幂组合而成。如果某一项 v_i 的指数小于 b_i , 则式 (1.1.11) 的求导运算将使该项为零。反之, 如果某一项 v_i 的指数大于 b_i , 则经过求导运算后代入 $v_i=0$, 也使该项变为零。因此, 在 $\phi(\mathbf{v})$ 中对混合矩有贡献的只是包含 $v_1^{b_1}, v_2^{b_2}, \dots, v_N^{b_N}$ 的项。由于当 B 为奇数时, 混合矩为零, 故当 B 为偶数时

$$E[x_1^{b_1} \cdots x_N^{b_N}] = \frac{1}{2^{B/2} (B/2)!} \cdot \frac{\partial^B}{\partial v_1^{b_1} \cdots \partial v_N^{b_N}} \left[\sum_{i=1}^N \sum_{k=1}^N v_i c_{ik} v_k \right]^{B/2} \Big|_{v_1=\dots=v_N=0} \quad (1.1.13)$$

在实际中应用较多的是 $N=4, b_i=1$ 的情况。此时

$$E[x_1 x_2 x_3 x_4] = c_{12} c_{34} + c_{13} c_{24} + c_{14} c_{23}$$

或写成

$$\begin{aligned} E[x_1 x_2 x_3 x_4] &= E[x_1 x_2] E[x_3 x_4] \\ &\quad + E[x_1 x_3] E[x_2 x_4] + E[x_1 x_4] E[x_2 x_3] \end{aligned} \quad (1.1.14)$$

后面在讨论自相关函数的估计特性时要用到这一结果。

1.2 正交投影原理

在信号线性最佳滤波理论中, 正交投影定理起了非常重要的作用。下面就对该定理加以介绍。

从几何意义上来说, 随机变量 $\{x_1, \dots, x_N, y_1, \dots, y_M\}$ 构成了 $N+M$ 维线性空间。为了把该空间转换成 Hilbert 空间, 需要对随机变量 u, v 定义内积如下

$$\langle u, v \rangle = E[uv] \quad (1.2.1)$$

根据这一定义, 两个随机变量**正交**, 即 $\langle u, v \rangle = 0$, 就是指不相关。另外随机变量 u 的模由下式给出

$$\|u\| = \sqrt{\langle u, u \rangle} \quad (1.2.2)$$

而两个随机变量 u, v 之间的**距离**, 则定义为

$$\|u - v\| = \sqrt{\langle (u - v), (u - v) \rangle} = \sqrt{E[(u - v)^2]}$$

相互正交的随机变量可以构成正交基。考虑 M 个相互正交的随机变量 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$, 那么 $\mathbf{Y} = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M\}$ 便是 $N+M$ 维 Hilbert 空间的一个 M 维子空间。其中 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$ 构成了该子空间的一组正交基。这样一来, 任意随机变量 x 可被正交分解成

$$x = \hat{x} + e \quad (1.2.3)$$

其中 $\hat{x}$ 在空间 Y 中；而 e 和空间 Y 垂直。我们通常把 $\hat{x}$ 称为随机变量 x 在空间 Y 的正交投影。图 1.2.1 给出了随机变量 x 在空间 $Y = \{e_1, e_2\}$ 上正交投影的示意图。

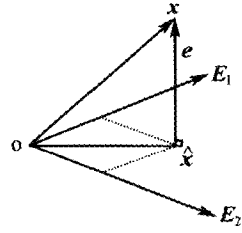


图 1.2.1 随机变量 x 在空间 $Y = \{e_1, e_2\}$ 上的正交投影

正交条件 $e \perp Y$ 意味着 e 和 Y 中的任意一个随机变量都正交，或 e 和 Y 的任意一个基都正交，即

$$\langle e, \varepsilon_k \rangle = E[e \varepsilon_k] = 0, \quad k = 1, 2, \dots, M \quad (1.2.4)$$

由于 $\hat{x}$ 在空间 Y 中，因此它可写为 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$ 的线性组合

$$\hat{x} = \sum_{k=1}^M a_k \varepsilon_k \quad (1.2.5)$$

为求出上式中的系数 a_k ，利用

$$\begin{aligned} \langle x, \varepsilon_k \rangle &= \langle \hat{x} + e, \varepsilon_k \rangle = \langle \hat{x}, \varepsilon_k \rangle + \langle e, \varepsilon_k \rangle = \langle \hat{x}, \varepsilon_k \rangle \\ &= \left\langle \sum_{i=1}^M a_i \varepsilon_i, \varepsilon_k \right\rangle = a_k E[\varepsilon_k^2] \end{aligned}$$

便可得到 $a_k = E[x \varepsilon_k] E[\varepsilon_k^2]^{-1}$ 。在其推导过程中，我们利用了正交条件 $\langle e, \varepsilon_k \rangle = 0$ 。把求得的系数代入式 (1.2.5) 后，有

$$\hat{x} = \sum_{k=1}^M E[x \varepsilon_k] E[\varepsilon_k^2]^{-1} \varepsilon_k \quad (1.2.6)$$

如果令矢量 $\varepsilon = [\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_M]^T$ ，则可以把上式写成更紧凑的形式。由于

$$E[x \varepsilon] = [E[x \varepsilon_1] \ E[x \varepsilon_2] \ \dots \ E[x \varepsilon_M]]^T$$

及

$$\Phi_{\varepsilon\varepsilon} = E[\varepsilon \varepsilon^T] = \text{diag}[E[\varepsilon_1^2] \ E[\varepsilon_2^2] \ \dots \ E[\varepsilon_M^2]]$$

其中 $\text{diag}[\]$ 表示对角线矩阵。这样式 (1.2.6) 便可写成

$$\hat{x} = E[x \varepsilon^T] E[\varepsilon \varepsilon^T]^{-1} \varepsilon \quad (1.2.7)$$

当我们把一个随机变量 x 看为线性空间的矢量时，有下述的正交投影原理。

正交投影原理 矢量 x 在其线性子空间 Y 上的正交投影 $\hat{x}$ 是 Y 中的一个矢量，它与 x

间的距离最小。

证明：令 y 为空间 Y 中的任一矢量，那么根据正交投影的性质，有 $(\hat{x} - y) \perp e$ ，其中 $e = x - \hat{x}$ ，故

$$\|x - y\| = \|\hat{x} - y + e\| = \|\hat{x} - y\| + \|e\|$$

因此只有当 $\|\hat{x} - y\| = 0$ ，即 $y = \hat{x}$ 时， y 与 x 间的距离为最小，其值为 $\|e\|$ 。

有时为了方便起见，我们把矢量 x 在其线性子空间 Y 上的正交投影写成 $\hat{E}[x|Y]$ 。

从滤波理论角度来看， x 在其线性子空间 Y 上的正交投影 $\hat{x}$ ，可以看成在空间 Y 中所能得到的关于 x 的最佳线性估计，这些我们将在后面的有关章节中讨论。

下面，我们介绍如何由线性子空间 Y 中的一组非正交基 $\{y_1, y_2, \dots, y_M\}$ 来得到其一组正交基。为此可令

$$\begin{aligned} \varepsilon_1 &= y_1 \\ \varepsilon_n &= y_n - \sum_{k=1}^{n-1} E[y_n \varepsilon_k] E[\varepsilon_k^2]^{-1} \varepsilon_k, \quad 2 \leq n \leq M \end{aligned} \quad (1.2.8)$$

根据正交投影原理，不难看出由上述方法所得到的 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$ ，为空间 Y 上的一组正交基。上述过程也被称为 Gram-Shmidt 正交化。如果定义子空间

$$\begin{aligned} Y_1 &= \{\varepsilon_1\} = \{y_1\} \\ Y_2 &= \{\varepsilon_1, \varepsilon_2\} = \{y_1, y_2\} \\ &\vdots \\ Y_n &= \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\} = \{y_1, y_2, \dots, y_n\} \end{aligned}$$

这里， $n = 1, 2, \dots, M$ 。那么式 (1.2.8) 右边的第二项就是空间 Y_n 中的矢量，即矢量 y_n 在空间 Y_n 上的正交投影。设该正交投影

$$\hat{y}_{n|n-1} = \sum_{k=1}^{n-1} E[y_n \varepsilon_k] E[\varepsilon_k^2]^{-1} \varepsilon_k \quad (1.2.9)$$

则

$$\varepsilon_n = y_n - \hat{y}_{n|n-1} \quad (1.2.10)$$

如果让

$$b_{nk} = E[y_n \varepsilon_k] E[\varepsilon_k^2]^{-1}, \quad 1 \leq k \leq n-1$$

及 $b_{nn} = 1$ ，那么由式 (1.2.8) 可以知道

$$\begin{aligned} y_1 &= \varepsilon_1 \\ y_n &= \varepsilon_n + \sum_{k=1}^{n-1} b_{nk} \varepsilon_k, \quad 2 \leq n \leq M \end{aligned}$$

或写成矩阵形式

$$\mathbf{y} = \mathbf{B}\boldsymbol{\varepsilon} \quad (1.2.11)$$

其中 $\mathbf{B}$ 是一个对角线的值为 1 的下三角矩阵, 或称为单位下三角矩阵; $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_M]^T$ 及 $\boldsymbol{\varepsilon} = [\varepsilon_1 \ \varepsilon_2 \ \cdots \ \varepsilon_M]^T$ 。

由上面的结果可以看出, 基 $y_1, y_2, \dots, y_M$ 和基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$ 所包含的信息是相等的, 即空间 $\mathbf{Y}$ 中的任一矢量都可以由它们表示出来。其不同之处就是基 $y_1, y_2, \dots, y_M$ 之间的相关性经正交化后被消除了。而变换前后的自相关矩阵有下述关系

$$\Phi_{yy} = E[\mathbf{y}\mathbf{y}^T] = \mathbf{B}E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T]\mathbf{B}^T = \mathbf{B}\Phi_{\varepsilon\varepsilon}\mathbf{B}^T$$

其中 $\Phi_{\varepsilon\varepsilon}$ 为单位矩阵。因此, 自相关矩阵 Φ_{yy} 被分解为单位三角矩阵和对角矩阵的乘积。另外利用式 (1.2.9) 容易证明, 矢量 $\mathbf{x}$ 在空间 $\mathbf{Y}$ 的正交投影不会随基的改变而变化, 即

$$\hat{\mathbf{x}} = E[\mathbf{x}\mathbf{y}^T]E[\mathbf{y}\mathbf{y}^T]^{-1}\mathbf{y} = E[\mathbf{x}\boldsymbol{\varepsilon}^T]E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T]^{-1}\boldsymbol{\varepsilon} \quad (1.2.12)$$

作为上式的一个应用, 我们可以把式 (1.2.9) 改写成

$$\hat{y}_{n|n-1} = E[y_n\mathbf{y}^T]E[\mathbf{y}\mathbf{y}^T]^{-1}\mathbf{y} \quad (1.2.13)$$

从而有

$$\varepsilon_n = y_n - \hat{y}_{n|n-1} = y_n - E[y_n\mathbf{y}^T]E[\mathbf{y}\mathbf{y}^T]^{-1}\mathbf{y}$$

上面的式子给出了直接由基 $y_1, y_2, \dots, y_M$ 构造正交基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$ 的方法。另外由正交投影原理可知, 式 (1.2.13) 所给出的 $\hat{y}_{n|n-1}$ 是由给定的 $y_1, y_2, \dots, y_{n-1}$ 所能够得到的关于 y_n 的最佳线性估计。

1.3 离散随机信号

一个实的离散随机信号定义为由随机变量组成的序列 $\{\dots, x(0), \dots, x(n), \dots\}$, 这里 $x(n)$ 表示位于时刻 n 处的一个随机变量。有时, 我们也把离散随机信号称为**随机序列**。

由上述定义可以看出, 离散随机信号在各个时刻的取值是不固定的。但是当某一离散信号作为随机序列的观测值时, 其在各个时刻的取值又都是确定的, 此时该信号被称为随机序列的一个**样本**。一般来说, 对随机信号的描述是十分复杂的, 它要求必须知道组成信号的各随机变量的所有的联合概率分布。图 1.3.1 给出了一随机序列示意图。

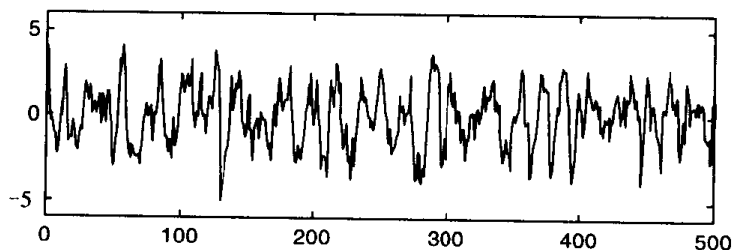


图 1.3.1 离散随机序列