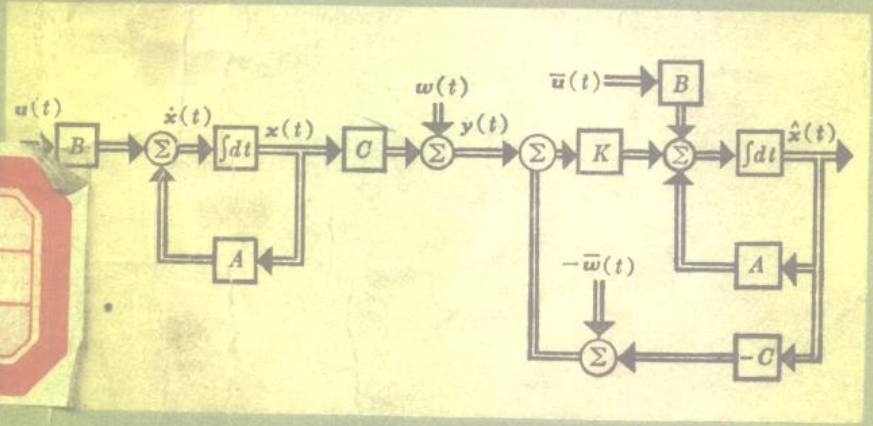
高等学校教学参考书

维纳与卡尔曼滤波理论导论

WEINA YU KAERMAN LÜBO LILUN DAOLUN

张有为 编



人民教育出版社

51.923
L88
.2

高等学校教学参考书

维纳与卡尔曼滤波理论导论

张有为 编

人民教育出版社

D67.3/20

内容简介

本书是介绍线性滤波理论基本内容的高等学校教学参考书，主要涉及维纳滤波器和卡尔曼滤波器。线性滤波理论是统计估计理论近代发展的一个重要分支，它成功地把统计的观点引入到通讯理论与控制理论，促进了通讯与控制学科的发展。目前这一理论可广泛应用于诸如雷达、通讯、导航、空间技术、自动控制、地质、地震、天体物理、生物物理、医学等学科。

全书共有六章，简要地叙述了滤波理论的发展史及其在统计估计理论中的地位、最小均方误差准则及维纳-霍夫方程、维纳滤波和预测、离散时间卡尔曼滤波、连续时间卡尔曼滤波、滤波理论的实际应用。为便于阅读和自学，书中的数学推导比较详尽。书末附有较详细的参考文献目录，以备深入了解这一领域的理论和应用的读者参考。

本书可供通讯与控制类专业的高年级大学生及有关科技工作者参考。

本书责任编辑　王忠民

高等学校教学参考书

维纳与卡尔曼滤波理论导论

张有为　编

*

人民教育出版社出版

新华书店北京发行所发行

北京印刷一厂印装

*

开本 787×1092 1/32　印张 8.75　字数 212,000

1980年7月第1版　1980年12月第1次印刷

印数　00,001—4,600

书号 15012·0247　定价 0.73 元

前　言

维纳(Norbert Wiener 1894～1964)富有创见的著作《平稳时间序列的外推、内插和平滑及其工程应用》公开发表至今整整三十年了[1]。维纳是著名的数学家，后被誉为信息理论家。他的重要贡献是将统计学的观点引入到通讯与控制理论中来，从统计本质上描绘出了通讯与控制问题的清晰的轮廓。维纳的著作不仅是一个崭新的创见，而且具有结合工程的实际意义，是线性滤波理论研究的一个重要的开端。

在自然科学中有两个重要的概念，就是能量与信息。能量这个概念在科学中所起的支配作用是人们所熟知的，而和能量的存在具有同样悠久历史的信息的概念，在科学上取得应有的地位，却是近三十年的事。三十年前信息论的开拓者香农(C. E. Shannon)所发表的经典性著作[2]标志着信息科学发展的起点。

从能量与信息的观点看，电工学在本世纪曾有过一次分裂[3]，德国人称其为强电流与弱电流之间的分裂，因而形成了动力工程学与通讯工程学。正是这种分裂使科学技术进入一个新的时代。

在第二次世界大战中，由于雷达的发明以及防空炮火控制的任务，把大量有修养的数学家和物理学家都动员到信息科学这个研究领域中来了。这时人们活跃于这个领域，并有许多重大的科学创造。数学家维纳对于控制论[3]和滤波理论[1]的研究成果，就是这时重大的科学创见之一。香农曾经说过："光荣应归于维纳教授，他关于平稳总体的滤波和预测问题的优雅的解答，在这个领域里，对我的思想有重大的影响"。

通讯与控制中的滤波问题，指的是从获得的信号与干扰中尽

可能地滤除干扰，分离出所期望的消息；或者说，是通过对一系列带有误差的实际测量数据的处理，得出所期望数据的估计值。

在第二次世界大战前后，通讯和控制都在自己的领域里得到了迅速的发展，分别建立了自己的理论。但是，抓住通讯与控制系统的共同特点，站在一个更概括的理论高度，来揭示它们的共同本质，显然是更为必要的。通讯与控制的共同特点的关键并不是环绕着电工技术，而是环绕着更为基本的信息的概念上。不论这信息是电的、机械的或是以其它方式传递的。也就是说，通讯与控制的共同特点在于都包含有一个信息的变换的过程。信息在时间上的分布是可测事件的离散或连续的时间序列，它带有某种随机的性质。通讯与控制系统本身的工程结构，也就必须适应它所接收和加工的信息的这种统计的性质。滤波理论就是站在这种认识的高度，来描述通讯与控制中消息过滤问题的本质的。

约在本世纪四十年代以前，要说明上述的观点是很困难的。因为用统计学的观点来说明通讯与控制问题的概念，对当时的传统思想来说，不但是新奇的，也是对传统思想的一种冲击。而今，这种概念已逐步为这个领域中的理论与技术工作者所熟悉。以致于使我担心本书所介绍的有些内容是否会被认为显得已经陈旧了。但是，不容置疑的是，在本世纪四十年代的最初几年中，维纳和柯罗莫格洛夫[4](A. H. Колмогоров)开始将统计的方法应用于通讯系统和控制系统的研究中是一种创举。使在滤波器设计中，具有经验性的甚至偶然性的过程，代之以具有科学判断的过程。

牛顿力学的辉煌成就，促进了统治着十七世纪至十九世纪自然科学思想的机械唯物论世界观的形成。机械唯物论者否认客观世界的偶然性，把偶然性与必然性绝对地对立起来。所以牛顿力学不能正确地反映微观世界，在微观世界面前束手无策。因此，滤波问题也就难于用基于牛顿力学的传统力学的方法来解决。所以，

维纳与柯罗莫格洛夫的工作，是富于创见的。

维纳的工作是从研究处在统计平衡的时间序列开始的，维纳证明：在一定条件下，处在统计平衡的时间序列的时间平均等于相平均。这是运用各态历经定理得到的一个重要的结论（见参考文献[3]第二章）。有了这个前提，就可以从时间序列的过去数据推知未来和预测未来。维纳正是基于这点提出了他的著名的滤波和预测理论的。滤波问题就是尽可能地恢复一个被噪声干扰了的消息流的问题。实质上，就是预测一个被噪声干扰了的时间序列的问题。因此，滤波问题也可以视为一个预测问题。从数学上讲，预测就是从一个时间序列的过去的数据来估算整个序列的统计参数。这种估算得出的是统计参数的平均值，它与客观实际是有一定的差距的，最佳预测应使这种误差为最小。维纳滤波所解决的是在最小均方误差准则下的滤波问题。

但是，维纳对他所提出的理论课题并没有全部完成。因为维纳滤波是通过求解维纳-霍夫方程得到最佳滤波器冲击响应的，而我们只能对于在平稳条件和有理谱密度情形下得出这个积分方程的解。而在一般情况下，由于求解维纳-霍夫积分方程存在有极大的困难，因而限制了它的应用与发展。正如维纳自己所说："这里发展的统计理论，要求我们对所观测的时间序列的过去具有充分的知识。但无论在什么场合，我们都不能满足这个要求，因为我们的观测不能追溯到无限的过去。为了超出这个范围，使我们的理论发展成为一个实用的统计理论，必须推广现有的采样方法"[3]。

维纳滤波理论的不足之处是显而易见的，在运用的过程中它必须把用到的全部数据存储起来，而且每一时刻都要通过对这些数据的运算才能得到所需要的各种量的估值。按照这种滤波方法设置的专用计算机的存储量与计算量必然极大，很难进行实时处理。虽经许多科技工作者的努力，在解决非平稳过程的滤波问题

维纳与柯罗莫格洛夫的工作，是富于创见的。

维纳的工作是从研究处在统计平衡的时间序列开始的，维纳证明：在一定条件下，处在统计平衡的时间序列的时间平均等于相平均。这是运用各态历经定理得到的一个重要的结论（见参考文献[3]第二章）。有了这个前提，就可以从时间序列的过去数据推知未来和预测未来。维纳正是基于这点提出了他的著名的滤波和预测理论的。滤波问题就是尽可能地恢复一个被噪声干扰了的消息流的问题。实质上，就是预测一个被噪声干扰了的时间序列的问题。因此，滤波问题也可以视为一个预测问题。从数学上讲，预测就是从一个时间序列的过去的数据来估算整个序列的统计参数。这种估算得出的是统计参数的平均值，它与客观实际是有一定的差距的，最佳预测应使这种误差为最小。维纳滤波所解决的是在最小均方误差准则下的滤波问题。

但是，维纳对他所提出的理论课题并没有全部完成。因为维纳滤波是通过求解维纳-霍夫方程得到最佳滤波器冲击响应的，而我们只能对于在平稳条件和有理谱密度情形下得出这个积分方程的解。而在一般情况下，由于求解维纳-霍夫积分方程存在有极大的困难，因而限制了它的应用与发展。正如维纳自己所说："这里发展的统计理论，要求我们对所观测的时间序列的过去具有充分的知识。但无论在什么场合，我们都不能满足这个要求，因为我们的观测不能追溯到无限的过去。为了超出这个范围，使我们的理论发展成为一个实用的统计理论，必须推广现有的采样方法"[3]。

维纳滤波理论的不足之处是显而易见的，在运用的过程中它必须把用到的全部数据存储起来，而且每一时刻都要通过对这些数据的运算才能得到所需要的各种量的估值。按照这种滤波方法设置的专用计算机的存储量与计算量必然极大，很难进行实时处理。虽经许多科技工作者的努力，在解决非平稳过程的滤波问题

简化理论分析，非线性滤波问题在理论处理上比线性滤波问题要困难和复杂的多。但是后来证明[7]：在一定条件下，在最小均方误差准则下得到的最佳线性系统是所有系统中的最佳者。

近代滤波理论的发展对于信息科学的发展是有重大贡献的，它概括了通讯与控制中信息过滤的统计本质。基础信息论虽然与滤波理论有同样长久的历史，但是它的发展却不如滤波理论这样活跃。这是由于滤波理论与通讯和控制中的许多课题有密切的联系，从而赋于了滤波理论以极大的生命力。滤波理论本来是一个小的研究领域，但是它联系着许多大的广泛的研究领域，因此它的价值已经超出了它起源时自身的价值。这也就是它能够继续活跃地向前发展的保证。

目前，滤波理论的研究已经越过了对于维纳滤波与卡尔曼滤波本身的研究阶段，许多科学工作者正在寻求对于非线性滤波的正确而严格的论述，寻求象解决线性滤波一样的统一的而不是个别的方法。这是对滤波理论问题研究的一种自然的延伸。这种发展是必然的，因为许多通讯中的调制方法以及大多数现实的火箭制导与控制问题本来就具有非线性的消息模型，因此必然导致发展非线性滤波的算法。虽然对于非线性滤波问题可以从修正的富克-普朗克 (Fokker-Planck) 方程得到非线性条件均值估计的精确表达式，但在运用上却是十分困难的。目前工程上使用的方法都是近似的，因此得到的结果还不是严格的最佳，甚至不能从根本上来评价各种方法的优劣，因为还没有得到衡量这种最佳的准则。就非线性滤波来说，目前的情况类似于维纳滤波在五十年代中期的情况，那时维纳滤波的发展似乎趋于停滞，许多滤波问题需要解决，因而向维纳滤波提出了新的挑战。当时卡尔曼和布西脱出了维纳滤波的困境，走向了一个新的有成效的发展方向。今天卡尔曼滤波所面临的形势，类似于五十年代中期维纳滤波所处的状况。

值得深思的是：沿着成功的卡尔曼滤波的方向寻求非线性滤波的严格解答能够成功吗？还是有新的方法而没有被发现呢？

在一个相当长的时间里，信息理论工作者和通讯科技工作者更熟悉的是关于消息与干扰的协方差信息，即维纳滤波问题；而控制科学工作者所更关心的是用状态空间模型描述消息与干扰，即卡尔曼滤波问题。遗憾的是，通讯科技工作者不那么关注卡尔曼滤波理论，而控制科技工作者又仅仅通过卡尔曼滤波学点估计理论，使得由于理论基础狭窄而蒙受损害，不能发挥应用的潜力。不少著名科学家都曾指出，这种差别应当逐步消失。从更概括的理论高度来观察通讯与控制问题显然是十分必要的。

三十年来滤波理论已经发展成了一个广阔的研究领域，可以有许多不同的方法来介绍它的内容，也可以选择不同的重点。本书是关于滤波理论方面导论性的参考书，主要介绍维纳滤波与卡尔曼滤波的基本概念，以及通过关于滤波理论发展史的介绍说明方法论上的演变。相对于滤波理论的丰富内容来说，本书也只能算做一个简介了。这样做的结果可能会遗漏许多重要内容，这对于某些科学家的成果来说也许是不公正的。然而由于编者知识的限制以及本书篇幅的限制，这似乎是不可避免的。

由于滤波理论是描述通讯与控制中信息提取的统计本质的，又由于卡尔曼滤波采用了状态变量法，因此本书必然涉及随机过程理论与矩阵等数学知识。如果读者对概率论、随机过程理论与矩阵过去毫无接触，那是很难顺利地接受本书各章的内容的。读者可自行选择有关书籍阅读，以弥补数学知识的不足。

由于编者的知识水平所限，难免有错误，请读者批评指正。北京航空学院毛士艺副教授对全稿提出了不少订正意见，谨表谢忱。

编者

1979 年 12 月

目　录

前　言……i

第一章　绪论……1

1-1　时间序列的滤波、预测与平滑……1

1-2　滤波理论的历史回顾……11

1-3　滤波理论在估计理论中的地位……28

第二章　最小均方误差准则及维纳-霍夫方程……38

2-1　最小均方误差准则……38

2-2　维纳-霍夫方程……50

2-3　求解维纳-霍夫方程的频谱因式分解法……57

2-4　求解维纳-霍夫方程的伯德-香农法……76

第三章　维纳滤波、预测与平滑……85

3-1　最佳滤波……85

3-2　滤波、预测与平滑……95

3-3　最佳纯预测器……107

3-4　最佳反馈滤波器……118

3-5　维纳滤波误差及其下限……123

3-6　典型消息谱的滤波误差……135

第四章　离散时间卡尔曼滤波……141

4-1　最小均方误差估计……141

4-2　卡尔曼滤波器的构成……151

4-3　正交投影定理与卡尔曼滤波公式导出……163

4-4　卡尔曼滤波器的性质及性能……175

第五章　连续时间卡尔曼滤波……183

5-1　连续时间卡尔曼滤波公式导出……183

5-2　黎卡蒂矩阵微分方程式的解……191

5-3　维纳-霍夫-卡尔曼积分方程……196

5-4　时不变卡尔曼滤波……211

第六章　滤波理论的实际应用……226
6-1　雷达目标状态估计……227
6-2　宇宙飞船轨道估计……243
6-3　随机信号最佳检测……246
6-4　调幅、调频、调相最佳接收……250
6-5　最佳控制与卡尔曼滤波器的结合……259
参考文献……264
编后记……270

第一章 绪 论

从数学的观点来说滤波理论是统计学中估计理论的一个重要分支；从工程的观点来看它又是系统工程研究的一个重要组成部分。滤波理论从一个新的高度描述了通讯与控制问题的统计本质，使我们对通讯与控制问题的认识更清晰更深刻。

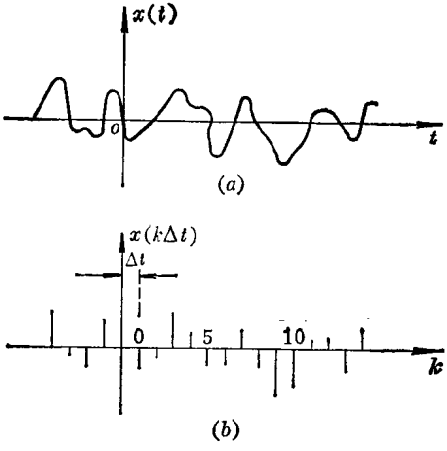
本章将通过关于时间序列的滤波、预测和平滑的介绍建立一个关于滤波的初步概念，并对时间序列的随机特性给以简要的数学描述。本章还将通过滤波理论的历史回顾，说明滤波理论的发展过程，不仅能为读者描绘出一幅基于最小均方误差估计的线性滤波理论的轮廓，而更重要的是希望读者了解在解决问题的方法上的变革。在叙述的过程中我们将提到诸如维纳、柯罗莫格洛夫、卡尔曼等科学家的名字，他们都是这个时期的众多科学家的代表。在本章的最后，介绍滤波理论在估计理论中的地位，以便使读者明了滤波理论与估计理论的关系。为此，我们还对估计理论中的经典概念做了极扼要的说明。读者在读这一章时，或许会感到有些内容难于理解，甚至难于接受，这是不奇怪的，因为关于滤波理论的详细讨论还是以后各章的事。这些是不影响我们建立起一个关于滤波的总的概念的。

1-1 时间序列的滤波、预测与平滑

时间序列 滤波、预测与平滑都是对时间序列而言的。所谓时间序列就是随机变量关于时间的函数。要对随机这个词下一个精确的定义，用来解释它的含意，似乎是不可能的。但是我们可以

1109328

通过具体的例子来表示它的意义。最著名的例子可以说是布朗运动了。布朗运动是英国植物学家布朗 (Robert Brown 1773～1858)在 1827 年首先发现的。布朗运动是悬浮在液体中的微粒(直径约为 10^{-3} 毫米) 所做的永不停止的运动。布朗在显微镜的视野里观察到了这种运动。微粒的运动是受来自各方液体分子的撞击所引起的。这就揭示了微观世界的秘密。类似的例子在通讯与控制中是到处可见的。雷达中的热噪声与目标噪声，飞行器遇到的湍流，心电图中的杂音尖峰，都是明显的例子。时间序列是这类物理现象的一个总称，它在时间坐标上的连续与离散分布如图 1-1所示。离散分布是连续时间函数以某一时间间隔采样的结果，称每个采样值为一个样本。相反，连续分布也可以认为是当采样时间间隔趋于无限小时离散分布的极限情况。



(a) 连续时间序列；　　(b) 离散时间序列。

图 1-1　时间序列

时间序列是对随机变量长期观察的结果，但随机变量的个别观察结果总是呈现为不规则的行为，要预测个别的结果几乎是不

可能的。但是，对它的长时间的观察的平均结果却表现出显著的规则性，统计学所描述的正是这种规则性的数学模型。

随机变量的分布函数(概率分布或概率密度)描述了随机变量的最完整的性质。可是在一系列的实际情形中，往往只要知道概率分布的某些数字特征就足够了，其中我们最关心的是均值与方差。

随机变量 x 的均值(又称为数学期望)，定义为

$$\mathrm{E}x=\int_{-\infty}^{\infty}xp(x)dx=m \tag{1-1}$$

式中 $p(x)$ 为随机变量 x 的概率密度。均值的物理意义是明显的。如果 x 是电压，那么 $\mathrm{E}x$ 就是这个电压的恒定分量；如果 x 是电流，那么 $\mathrm{E}x$ 就是这个电流的直流分量。

随机变量 x 的平方均值定义为

$$\mathrm{E}x^2=\int_{-\infty}^{\infty}x^2p(x)dx \tag{1-2}$$

它也有明显的物理意义，如果随机变量 x 是电压或电流，那么 $\mathrm{E}x^2$ 就是这个电压或电流分布在 1 欧姆电阻上的平均功率。这个功率是由随机电压(或电流)的恒定分量(或直流)和起伏分量所引起的。

随机变量 x 的方差定义为

$$\mathrm{Var}\,x=\mathrm{E}[x-\mathrm{E}x]^2=\mathrm{E}[x-m]^2=\sigma^2 \tag{1-3}$$

方差的物理意义在工程领域是很重要的。如果随机变量 x 是电压或电流，那么 σ^2 在数量上等于起伏分量在 1 欧姆电阻上消耗的平均功率。

均值、平方均值及方差是有联系的

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

$$=\int_{-\infty}^{\infty}x^2p(x)dx-2m\int_{-\infty}^{\infty}xp(x)dx+m^2\int_{-\infty}^{\infty}p(x)dx$$

$$=Ex^2-2m^2+m^2$$

$$=Ex^2-m^2 \tag{1-4}$$

或记为

$$\sigma^2=Ex^2-(Ex)^2 \tag{1-5}$$

这意味着，如果随机变量 x 是电压或电流，方差就是在 1 欧姆电阻上消耗的总平均功率与功率的恒定分量之差。方差表征了随机变量的一切值在其平均值的两旁的分布情况，即可用方差来度量离散或密集的情况。除了选择方差作为离散的度量之外，还有其它方法[8]。我们特别采用方差作为离散度量的原因，不是出于任何逻辑上的必要性，而是出于习惯和运算法则简单，而其它方法都没有这样简单的特性。例如方差的加法规则就是最简单的

$$\sigma_\Sigma^2=\sigma_1^2+\sigma_2^2+\cdots+\sigma_n^2 \tag{1-6}$$

几个独立随机变量和的方差等于每个独立变量方差的和。

在随机变量的许多典型分布中，人们似乎对正态分布（或称高斯分布）有些偏爱。有些人认为：只要能够进行大量的足够的精确的观察，那么作为一种理想的极限形式，几乎一切的统计分布都趋于正态分布。显然这种说法太夸张了，应该加以纠正。但是，不能否认在大量的应用中，我们遇到的许多分布至少是近似正态的。关于这点许多人是相信的，实验工作者相信它，因为它是一个数学定理；而数学家相信它，认为它是一个实验的事实。数学的证明告诉我们，在某些限制性条件下，期望一个正态分布是合理的；而统计的经验在事实上证明，分布常是近似正态的。

随机变量 x 的正态分布习惯上记为 $N(m,\sigma^2)$，其概率密度为

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2\sigma^2}(x-m)^2\right\} \tag{1-7}$$

其数学期望与方差为

$$\mathrm{E}x=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}x\exp\left\{-\frac{1}{2\sigma^2}(x-m)^2\right\}dx=m \tag{1-8}$$

$$\begin{aligned}\mathrm{Var}\ x&=\mathrm{E}[x-m]^2\\&=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}[x-m]^2\exp\left\{-\frac{1}{2\sigma^2}(x-m)^2\right\}dx=\sigma^2\end{aligned} \tag{1-9}$$

由此可见，一个正态分布由它的均值与方差完全确定了。这是人们乐于把实验的事实看作为近似正态分布的重要原因。因为一般的概率分布并不具有这个特性。

对于 n 维随机向量 $\boldsymbol{x}$，存在有均值向量和方差阵，可以类似于一维的情况来定义。均值向量为

$$\mathrm{E}\boldsymbol{x}=\begin{bmatrix}\mathrm{E}x_1\\\mathrm{E}x_2\\\vdots\\\mathrm{E}x_n\end{bmatrix}=\int_{-\infty}^{\infty}\cdots\int\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}p(x_1,x_2,\cdots,x_n)dx_1dx_2\cdots dx_n \tag{1-10}$$

或简写为

$$\mathrm{E}\boldsymbol{x}=\int_{-\infty}^{\infty}\boldsymbol{x}p(\boldsymbol{x})d\boldsymbol{x} \tag{1-11}$$

其方差阵定义为如下的 $n\times n$ 矩阵：

$$\mathrm{Var}\,\boldsymbol{x}=\mathrm{E}[\boldsymbol{x}-\mathrm{E}\boldsymbol{x}][\boldsymbol{x}-\mathrm{E}\boldsymbol{x}]^{\mathrm{T}} \tag{1-12}$$

可记为

$$\begin{aligned}\mathrm{Var}\,\boldsymbol{x}=&\int_{-\infty}^{\infty}\cdots\int[x_i-\mathrm{E}x_i][x_j-\mathrm{E}x_j]\\&\times p(x_1,\cdots,x_n)dx_1dx_2\cdots dx_n\end{aligned} \tag{1-13}$$

或简写为

$$\mathrm{Var}\,\boldsymbol{x}=\int_{-\infty}^{\infty}[\boldsymbol{x}-\mathrm{E}\boldsymbol{x}][\boldsymbol{x}-\mathrm{E}\boldsymbol{x}]^{\mathrm{T}}p(\boldsymbol{x})d\boldsymbol{x} \tag{1-14}$$

$\mathrm{Var}\,\boldsymbol{x}$ 是一个非负定阵。

在研究滤波问题时，我们很关心协方差信息。设 n 维随机向量 $\boldsymbol{x}$ 和 m 维随机向量 $\boldsymbol{y}$ 的联合概率密度为 $p(\boldsymbol{x},\boldsymbol{y})$，我们定义它们的协方差矩阵 $\mathrm{Cov}[\boldsymbol{x},\boldsymbol{y}]$ 为如下的 $n\times m$ 矩阵：

$$\begin{aligned}\mathrm{Cov}[\boldsymbol{x},\boldsymbol{y}]&=\mathrm{E}[\boldsymbol{x}-\mathrm{E}\boldsymbol{x}][\boldsymbol{y}-\mathrm{E}\boldsymbol{y}]^{\mathrm{T}}\\&=\int_{-\infty}^{\infty}[\boldsymbol{x}-\mathrm{E}\boldsymbol{x}][\boldsymbol{y}-\mathrm{E}\boldsymbol{y}]^{\mathrm{T}}p(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y}\end{aligned}\tag{1-15}$$

在一个随机向量的方差阵中，其主对角线元是各分量的方差，而其第 i 行第 j 列元($i\neq j$)则是第 i 分量和第 j 分量的协方差。

类似于一维的情形，协方差阵与均值向量间存在有如下的关系：

$$\mathrm{Cov}[\boldsymbol{x},\boldsymbol{y}]=\mathrm{E}\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}-(\mathrm{E}\boldsymbol{x})(\mathrm{E}\boldsymbol{y})^{\mathrm{T}}\tag{1-16}$$

若协方差阵 $\mathrm{Cov}[\boldsymbol{x},\boldsymbol{y}]$ 是为零的矩阵，则

$$\mathrm{E}\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}=(\mathrm{E}\boldsymbol{x})(\mathrm{E}\boldsymbol{y})^{\mathrm{T}}\tag{1-17}$$

即 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 不相关。两个独立的随机向量一定是不相关的，但一般不能说，两个不相关的随机向量是相互独立的[8]。

所谓时间序列就是随机变量关于时间的函数。作为滤波问题所考虑的对象，常常不是单个的随机变量或随机向量，而将涉及一串随机变量或随机向量。研究作为时间函数的随机变量，就导致我们来研究随机过程。所谓随机过程，就是依赖于时间的随机变量的全体。如果对试验的每一结果 ω 指定一个时间函数 $x(t,\omega)$，我们就建立了一个随机过程，通常简记为 $x(t)$。对于随机向量过程则应记为 $\boldsymbol{x}(t)$。图 1-2 是一个随机过程的记录。

随机过程的均值定义为

$$\mathrm{E}[x(t,\omega)]=\lim_{T_1,\,T_2\to\infty}\frac{1}{T_1+T_2}\int_{-T_1}^{T_2}x(t,\omega)dt=m(t)\tag{1-18}$$

随机过程 $x(t)$ 的时间 t 和 τ 的协方差函数为