0 绪论

在“信号与系统”与“离散时间信号处理”的基本课程中，已经讨论了连续和离散确定性信号的时域和频域表示方法，以及系统对信号所做的各种处理。但在自然界中遇到的实际信号大都具有随机特性，这样的信号称为随机信号。随机信号的表示、分析和处理与确定性信号有许多不同之处。确定性信号的分析工具是随机信号分析的重要基础，但随机信号分析和处理又有许多不同的方面。本书要讨论的主要问题是离散随机信号的分析和处理方法，主要包括如下几方面。

(1) 信号的特性

随机信号是一个随机过程，它可以是时间连续的，也可以是时间离散的，本书中主要讨论时间离散的随机信号。一个离散随机过程在每一个时刻对应一个随机变量，它服从一个概率分布函数，多个时刻对应的多个随机变量服从联合概率分布函数，这是随机过程的最基本的数学特征。在许多信号处理问题中，需要的是随机过程的更简单的特征，例如，对于各种统计平均量，最常用的是期望值和相关函数。

在对实际信号的分析中，一般只能得到随机过程的一次实现，随机过程的一次实现称为一个时间序列，通过一个时间序列分析或估计随机信号的参数和特征，以提取有用的信息。

当随机信号能够用更加特殊的模型表示，分析方法往往会更加有效。例如，最常用的一种模型是线性系统模型，其随机过程由一个差分方程来刻画，差分方程的驱动信号是一个特性已知的简单的随机过程；另一种常用的信号模型是谐波模型，其随机信号由一组具有随机相位的正弦信号组成，噪声干扰就嵌在这个信号模型中。当一个时间序列符合这样的特殊模型时，许多分析和处理方法变得简单甚至具有简捷的闭式方程。

随机信号处理中的一个重要问题，即功率谱估计的现代方法，就是建立在这样一些模型的基础上的。

（2）统计意义下的最优滤波

随机信号处理的一个重要问题是波形估计，这类问题也称为滤波问题。滤波一词的含义已远不再限于“选频”这样一个狭窄的概念，凡是对时间序列进行处理以改善对信号的理解的过程均称为滤波。在随机信号作为输入和输出的滤波器中，滤波器的输出怎样是最优的？这个最优的评价准则应该考虑随机信号的特点和不同的应用对象。对随机信号，一个最常用的评价准则是均方误差最小准则。在平稳条件下，对一个期望波形的最优估计问题归结为维纳（Wiener）滤波器理论。作为线性滤波器理论上的目标，Wiener 滤波器给出了实际算法可达到的按均方误差准则的最优解，如果按照最大信噪比准则，则可以得到匹配滤波器作为最优解。在非平稳条件下，一个递推的最优滤波器是卡尔曼（Kalman）滤波，Kalman 滤波的目标是对随机动力系统的状态变量进行估计和预测，递推结构使得 Kalman 滤波可以方便地跟踪非平稳变化的状态变量，并可以扩展到非线性系统。

（3）具备“学习能力”的实际滤波算法

Wiener 滤波器给出了线性滤波在统计意义下的最优解，但 Wiener 滤波器的实现需要对输入信号的先验统计，并要求输入信号的二阶统计特征是已知的。这在实际应用中是不现实的，一种实际的实现方法是通过学习（或者说是一种递推的调节算法）自适应地获得这些统计量，以得到实际上可实现的滤波器。初始时滤波器的系数是预置的，可能产生较差的估计结果，但随着对环境的学习递推地对滤波器系数作调节，估计结果将收敛到逼近于最优滤波器（Wiener 滤波器）的性能。

这类具有学习能力的线性滤波器称为线性自适应滤波器，它已经获得广泛应用，如通信信道的自适应均衡、线性系统辨识、噪音对消、波束形成等。

这类线性自适应滤波器需要一个“教师”的指导，这个“教师”就是一个期望响应信号。例如，在通信系统的自适应均衡器的设计中，通过通信系统接通时的训练序列来提供期望响应。通过一些非线性的调节算法，也可以在无“教师”的条件下逐渐改善滤波器性能，构成信号的盲处理系统，最常用的是盲反卷和盲分离。

（4）更多信息的利用和挖掘

在线性自适应滤波和功率谱估计的应用中，仅使用了随机信号的二阶矩。二阶矩对高斯过程给予完全的表示，但对于非高斯过程，二阶矩是不完整的，通过高阶矩可以获得随机信号的更多有用信息。例如二阶矩和功率谱是相位盲的，它们不包含相位信息，利用这些特征对系统进行辨识时，只能处理最小相位系统。利用高阶矩和多谱可以提取信号的相位信息、实现非最小相位系统的辨识，分离高斯和非高斯分量，以及实现系统的盲反卷等。

（5）对时间（空间）-频率关系的局域性分析

人们的注意力容易被信号中的瞬变或景物中的运动所吸引，而容易忽略其中的平稳

环境。这些瞬变和运动包含了更重要的信息，同时，对于随机信号，这些局域的瞬变过程具有非平稳统计特性。为了获取这些重要信息，将分析工具局域化到瞬变时刻的附近，将得到更精细的结果，因此具有时频局域性的分析工具是重要的。

傅里叶变换是线性系统分析最重要的工具，要得到一个信号 $f(t)$ 的傅里叶变换，必须在整个时间轴上对 $f(t)$ 和 $e^{i\omega t}$ 进行混合，因此傅里叶变换没有能力抽取一个信号的局域性质，即它没有能力抽取或定位信号在某个时间附近的瞬变特性。

因此，傅里叶变换对瞬态或非平稳信号的局域特性分析是无能为力的。人们研究了各种时-频分析方法，目的是进行局域的时间域和频率域的分析，典型的线性时-频分析工具是短时傅里叶变换及离散格伯（Gabor）展开，而维格纳分布则属于非线性时频分析方法。小波变换属于这一重要领域，它是一种线性时-频分析工具，但不管是连续域还是离散域，小波变换都有简洁的正变换和反变换关系，并具有快速的离散小波变换算法，它既可以作为信号分析的工具，又可以作为信号处理的工具，因此得到了更广泛的应用。

（6）应用举例

举例是理解各种算法的最有帮助的方式，书中列举了许多例子，既有为理解原理的说明性的例子，也有许多数字仿真实例。

现代信号处理的应用非常广泛，难以一一列举，下面仅简要介绍几个常用的应用领域。

（1）噪声消除

降低噪声的影响可能是信号处理中最常见的任务，也是一个困难的任务。我们每个人在生活中都有这样的经历，在一瓶纯净水中滴入几滴墨汁是很容易的，但要去除这些墨汁的影响却要难得多。噪声消除与此很相似，在信号的采集、传输、存储过程中，都难免被噪声污染，要得到尽可能纯净的信号波形，有必要消除噪声。实际上要完全去除噪声几乎是不可能的，只要将噪声影响降低到不影响我们对信号的使用，也就达到了目的。噪声消除的方法很多，对具体问题可以选择一种或几种方法的组合。例如，从原理上，对于平稳随机信号和噪声，Wiener 滤波器是均方误差最小意义上的最优线性滤波器，即 Wiener 滤波器的输出与纯净信号误差的均方值是最小的；而自适应滤波器是一种更易于实现的装置，它通过“学习”过程，达到或逼近 Wiener 滤波器的性能。Kalman 滤波器也可以用于消除噪声，当对一个动力系统的状态的测量值混入了噪声时，Kalman 滤波器用于估计状态的真实波形，Kalman 滤波可以应用到非平稳环境，通过扩展，还可以应用于非线性系统。利用小波变换或时-频分析，通过门限方法，也可以有效地消除噪声；高阶统计量分析中，利用高斯过程的高阶累积量为零的性质，可以有效地提取非高斯信号。

（2）预测编码

在语音编码和图像编码中，预测编码技术都是一类重要的编码方法。在预测编码技术中通过设计最优的线性预测器，由信号的过去值尽可能好地预测当前值，并得到小的预

测误差。在信源编码中，将对原始信号序列的编码改换成对预测误差序列的编码。对于最优预测器，信号的预测误差序列是低能量的近似白噪声序列，通过量化后（如果是无失真编码，则不经过量化过程）可以用比较简单的熵编码方法进行有效的编码。在信号处理中，最优线性预测器可以看作是 Wiener 滤波器的特例，并且可以用莱文森-杜宾(Levinson-Durbin)算法快速求解。

(3) 信道均衡

在通信系统中，发射方传送的信号序列，经过信道的畸变和噪声的干扰，在接收端接收到的信号已经很恶劣，直接用检测器进行检测会产生高的误码率。因此，通信系统在接收端设置一个均衡器用于补偿信道的畸变和噪声干扰。理论上，这个均衡器是一个最优滤波器，目标是使滤波器的输出最逼近于发射端发送的信号序列。实际中，经常采用“训练序列”设计自适应均衡器。在通信系统接通时，发射端传送一个“约定的序列”给接收端，接收端可以重新产生这个“约定序列”即训练序列，训练一个自适应滤波器，将接收的已畸变和被噪声干扰的序列转化为尽可能逼近于训练序列的序列，这实际是在设计一个对信道的逆系统。信道均衡器可归结于系统求逆的应用范围，这是自适应滤波的一个重要应用方向。

(4) 延迟估计

在雷达和声呐系统中，需要估计发射波和反射波之间的延迟时间。在信号处理中，通过互相关分析和互功率谱进行延迟估计是最基本的方法。利用高阶统计量，由于避免了高斯噪声的影响，因此可能获得更可靠的延迟估计。

(5) 波束形成和入射波方向估计

在雷达、声呐和移动通信系统中，经常采用在空间分布的天线阵列进行信号接收和处理。例如，接收和加强来自一个预定角度的波，而抑制其他方向的波，这称为波束形成问题。天线阵列的另一个应用是，估计各个入射波的方向角度。在同一时刻，天线阵列接收的信号按空间分布，是空域信号，本书重点研究按时间变化的信号，即时间信号处理。需要注意，空间信号处理与时间信号处理非常相似，时间信号处理的算法稍加变化或不作变化即可应用于空间阵列信号处理中，如用 Wiener 滤波器解决波束形成问题。在第 5 章，讨论了利用信号的自相关矩阵的特征空间特性估计时间信号的频率，导出了性能良好的 MUSIC 和 ESPRIT 算法，这些算法应用于空间天线阵列信号中，对应的是估计各个入射波的方向。

1 随机信号基础及模型

本章讨论全书用到的一些基础知识。首先讨论随机信号的表示和一些基本特征量的定义，重点是平稳随机信号的自相关函数和功率谱及其通过线性系统后的表达式，这些内容实际是对随机过程的复习，但这里重点讨论的是离散情况，而随机过程的数学课程更注重于连续情况。本章还讨论了随机信号的常用模型，包括谐波模型和有理分式模型，一方面，这些模型为后续章节提供了随机信号的实例，另一方面，也是更重要的，这些模型是随机信号功率谱密度估计的现代方法的主要基础。

1.1 随机信号基础

在实际中遇到的大多数信号是具有随机性质的信号，该随机信号可以用随机过程来表示。因此，本书中随机过程和随机信号是等同的名词。随机信号在一个确定时刻的值是一个随机变量，它服从概率分布函数，在绝大多数情况下它也存在一个概率密度函数。随机信号在不同时刻的取值构成随机矢量，它服从联合概率分布函数和存在联合概率密度函数。

对于一个随机信号，进行多次实验记录的信号波形可能都是不同的，每次实验的波形称为随机信号的一次实现，所有实现的集合构成这个随机过程。因此，对于一个随机信号，它既有不确定性，也具有确定性的规律。随机信号每一次实现的波形是不同的，一般情况下，我们无法准确预测当前实现中随机信号在某一时刻的取值，但是，一个随机信号服从确定的概率分布和联合概率分布；对于一个指定的时刻，信号取值的概率分布和概率密度函数是确定的，对于一组指定的

时间集合,联合概率分布和联合概率密度函数也是确定的;另外,对于一个随机信号,它具有一些确定的统计特征量(简称统计量),这些统计特征量反映了信号的许多本质的性质,通过对这些统计量的分析,可以获得随机信号中许多重要的信息。

图 1.1 是一个随机信号的三次不同的实验结果。这个随机信号有一个初相位是随机变量的正弦波被噪声所干扰,由于正弦信号的初相位是随机的,因此每次实验前无法预测这个初相位,每个时刻的干扰噪声都服从相同的正态(高斯)分布,但取值是不可能准确预测的。尽管如此,我们也可以从这些实验的数据中估计出正弦波形的频率和幅度等重要信息。

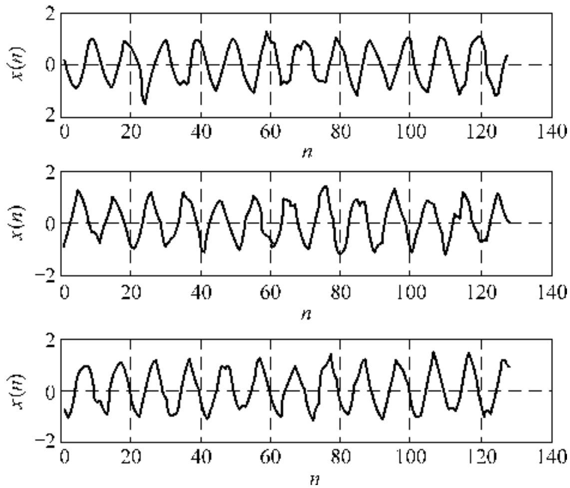

图 1.1　被噪声干扰的初相位是随机值的正弦波

物理世界的大多数随机信号是连续信号,为了便于处理,需要通过采样得到离散随机信号。本书主要针对离散随机信号展开讨论,因此后文中如果不特殊说明,均指离散随机信号。

随机信号的各种统计特征量反映了随机信号的总体性质,一般需要通过随机信号的联合概率密度函数来求得这些统计特征量,它们就是反映随机信号所有实现的集合的性质,称为随机信号的汇集特征或汇集平均。在工程实际中,可以获得的往往只有随机信号的一次实现,统计信号处理中的一个重要任务就是通过这样一次实现,去估计随机信号的各种统计特征量及其这些统计量的各种变换形式,从而提取有用信息。我们称离散随机信号的一次实现为一个时间序列,记为$\{x(n)\}$,或简单记为 $x(n)$。$x(n)$只是表示随机信号的一次实现,且只是一个时间序列,而随机信号是它的所有可能实现的集合。一个随机过程可以用$\{x(n,\xi)\}$表示,ξ代表随机过程的一次实验,$\{x(n,\xi)\}$表示所有实验的集合。但是为了叙述简单,本书也用符号 $x(n)$表示一个随机信号,这样做一般并不会引起歧义。

1.1.1 随机过程的基本特征

为简单起见，用一个时间序列表示一个离散时间随机信号，时间序列是由一组按时间顺序排列的观测值所组成的。一个离散随机信号可以用如下时间序列表示：

$$\cdots,x(-1),x(0),x(1),\cdots,x(n-1),x(n),\cdots$$

对于一个随机过程，比较完整的描述是对于任意取的时间集合，给出他们的联合概率分布函数，一般地用

$$F(x_1,x_2,\cdots,x_M,n_1,n_2,\cdots,n_M)=P\{x(n_1)\leqslant x_1,x(n_2)\leqslant x_2,\cdots,x(n_M)\leqslant x_M\}$$

表示随机信号分别在$n_1,n_2,\cdots,n_M$时刻取值的联合概率分布函数，这里 $P\{\xi\}$表示事件ξ发生的概率。如果$F(x_1,x_2,\cdots,x_M,n_1,n_2,\cdots,n_M)$分别对 $x_1,x_2,\cdots,x_M$是可导的，定义

$$p(x_1,x_2,\cdots,x_M,n_1,n_2,\cdots,n_M)=\frac{\partial F(x_1,x_2,\cdots,x_M,n_1,n_2,\cdots,n_M)}{\partial x_1\partial x_2\cdots\partial x_M}$$

为随机信号在$n_1,n_2,\cdots,n_M$时刻取值的联合概率密度函数。在信号处理的文献中，通过引入冲激函数 $\delta(x)$，可以将取值连续和取值离散的随机信号的联合概率密度函数作统一处理。当

$$p(x_1,x_2,\cdots,x_M,n_1,n_2,\cdots,n_M)=p(x_1,n_1)p(x_2,n_2)\cdots p(x_M,n_M)$$

时，称随机信号在这些时刻的取值是互相统计独立的。如果

$$p(x_1,x_2,\cdots,x_M,n_1,n_2,\cdots,n_M)=p(x_1,x_2,\cdots,x_l,n_1,n_2,\cdots,n_l)p(x_{l+1},\cdots,x_M,n_{l+1},\cdots,n_M)$$

则称该随机信号在时间子集$n_1,n_2,\cdots,n_l$的取值和时间子集$n_{l+1},n_{l+2},\cdots,n_M$的取值是统计独立的，但在每个时间子集内不一定独立。

在信号处理中，联合概率密度函数更多的是理论上的意义，实际中比较难获取。用联合概率密度函数，可以定义随机信号的一些常用的统计特征量，这些统计特征量可能会通过记录的一个随机信号的一次实现进行近似估计。随机信号最常用的统计特征量是它的1阶和2阶特征，包括均值、相关和协方差函数，它们的定义分别如下。

(1) 均值(1阶特征，或称1阶矩)：

$$\mu(n)=E[x(n)]=\int xp(x,n)\mathrm{d}x \tag{1.1}$$

(2) 自相关(2阶特征，或称2阶矩)：

$$r(n_1,n_2)=E[x(n_1)x^*(n_2)]=\iint x_{n_1}x_{n_2}p(x_{n_1},x_{n_2},n_1,n_2)\mathrm{d}x_{n_1}\mathrm{d}x_{n_2} \tag{1.2}$$

(3) 自协方差函数(2阶特征，或称2阶中心矩)：

$$\begin{aligned}c(n_1,n_2)&=E[(x(n_1)-\mu(n_1))(x(n_2)-\mu(n_2))^*]\\&=r(n_1,n_2)-\mu(n_1)\mu^*(n_2)\end{aligned} \tag{1.3}$$

(1.1)式中，$p(x,n)$表示随机信号在 n 时刻取值的概率密度函数；(1.2)式中，$p(x_{n_1},x_{n_2},n_1,n_2)$表示随机信号分别在$n_1$和$n_2$时刻取值的联合概率密度函数。相似地还可以定

义高阶统计量(高阶矩),例如

$$E[x(n_1)x(n_2)x(n_3)x(n_4)]$$
$$=\iiiint x_{n_1}x_{n_2}x_{n_3}x_{n_4}p(x_{n_1},x_{n_2},x_{n_3},x_{n_4},n_1,n_2,n_3,n_4)\mathrm{d}x_{n_1}\mathrm{d}x_{n_2}\mathrm{d}x_{n_3}\mathrm{d}x_{n_4}$$

是四阶矩的定义,高阶统计量的性质和应用在第 7 章详细讨论。

自相关函数表示随机信号在两个不同时刻取值的相关性,自协方差函数是去掉均值后的相关性。在信号处理中,经常将均值不为零的信号去均值后再进行处理,在本书中,如不加特别说明,总假设随机信号是零均值的,此时,自相关等于自协方差。对于一个实际信号,如果均值不为零,则通过先估计均值,然后将每个样本减去均值,从而使其变成零均值信号后再作处理。

例 1.1.1 设一个随机信号为

$$x(n)=A\cos(\omega n+\varphi)$$

其中 φ 是$[0,2\pi]$内均匀分布的随机变量;A 有两种情况:(1)A 是与 φ 独立的随机变量,服从均值为零、方差为σ^2的正态分布,(2)A 是常数。求 $x(n)$的均值和自相关函数。

解

$$\mu(n)=E[A\cos(\omega n+\varphi)]=E(A)E[\cos(\omega n+\varphi)]$$
$$=E(A)\int_0^{2\pi}\frac{1}{2\pi}\cos(\omega n+\varphi)\mathrm{d}\varphi=0$$
$$r(n_1,n_2)=E[A\cos(\omega n_1+\varphi)A\cos(\omega n_2+\varphi)]$$
$$=\frac{1}{2}E(A^2)E[\cos\omega(n_1-n_2)+\cos(\omega n_1+\omega n_2+2\varphi)]$$

由于

$$E[\cos(\omega n_1+\omega n_2+2\varphi)]=\frac{1}{2\pi}\int_0^{2\pi}\cos(\omega n_1+\omega n_2+2\varphi)\mathrm{d}\varphi=0$$

得到

$$r(n_1,n_2)=\frac{1}{2}E(A^2)\cos\omega(n_1-n_2)$$

对于情况(1),$E(A^2)=\sigma^2$,因此

$$r(n_1,n_2)=\frac{1}{2}\sigma^2\cos\omega(n_1-n_2)$$

对于情况(2),$E(A^2)=A^2$,得

$$r(n_1,n_2)=\frac{1}{2}A^2\cos\omega(n_1-n_2)$$

#

信号处理中最常用和分析方法最成熟的是对平稳随机过程的分析。平稳随机过程是指对任意 M 个不同时刻,信号取值的联合概率密度函数与起始时间无关,即$\{x(n_1),$

$x(n_2),\cdots,x(n_M)\}$和$\{x(n_1+k),x(n_2+k),\cdots,x(n_M+k)\}$的联合概率密度函数对任意整数 k 是相同的。实际中更常处理的是所谓宽平稳随机过程(WSS,也称为广义平稳)。宽平稳是指随机过程的各阶矩与起始参考时间无关,例如一个二阶 WSS 应满足

$$\begin{aligned}&\mu(n)=\mu\\&r(n_1,n_2)=r(n_1-n_2)=r(k),\quad k=n_1-n_2\end{aligned}\tag{1.4}$$

严格地讲,K 阶宽平稳过程是指 K 阶矩满足与起始参考时间无关的性质。但在实际中,针对不同的应用,我们只关心某阶平稳,如功率谱估计只关心 2 阶宽平稳。在后面的叙述中,我们隐含地认为宽平稳是指 2 阶宽平稳。因为宽平稳是更常用的,因此后面提到的平稳性,如果不加限定条件,指得就是宽平稳,一般平稳性用术语“严格平稳”来专指。宽平稳比严格平稳的条件要宽松,在严格平稳条件下,如果随机过程的均值和自相关存在,则一定是宽平稳的;反之,宽平稳情况下,推不出严格平稳性。

对于平稳信号,自相关的定义可以更简单地写为

$$r(k)=E[x(n)x^*(n-k)]$$

其中,$r(0)=E[|x(n)|^2]=\int_{-\infty}^{+\infty}|x|^2p(x)\mathrm{d}x$,表示信号的平均功率。对零均值信号,$r(0)=E[|x(n)|^2]=\sigma^2$,即自相关函数在零序号的值等于随机信号的方差。

注意,在今后使用自相关函数、均值和方差等特征量时,可以带下标,也可以不带下标,譬如要指明是随机信号 $x(n)$和 $y(n)$的自相关,分别用符号$r_x(k)$和 $r_y(k)$表示,如果不写下标,则表明自相关 $r(k)$所对应的随机信号是明确的,不必用下标区分。

例 1.1.2　针对例 1.1.1 所讨论的信号,有

$$\mu(n)=\mu=0$$

$$r(k)=r(n_1,n_2)=\frac{1}{2}E(A^2)\cos\omega(n_1-n_2)=\frac{1}{2}E(A^2)\cos\omega k$$

该信号是 2 阶宽平稳的随机信号。

♯

不难验证自相关函数有如下最常用性质:

(1) 原点值最大　$r(0)\geqslant|r(k)|$;

(2) 共轭对称性　$r(-k)=r^*(k)$,对于实信号有 $r(-k)=r(k)$,即实信号的自相关函数是对称的;

(3) 半正定性　对于任给的 a_i,满足 $\sum_i\sum_j a_ia_j^*r(i-j)\geqslant 0$。

对两个不同的随机过程,可以定义它们的互相关为

$$r_{xy}(n_1,n_2)=E[x(n_1)y^*(n_2)]$$

如果他们是联合宽平稳的,则它们的互相关为

$$r_{xy}(k)=E[x(n)y^*(n-k)]\tag{1.5}$$

互相关有如下性质：

$$r_{xy}(-k)=r_{yx}^{*}(k)$$

$$|r_{xy}(k)|^{2}\leqslant|r_{x}(0)||r_{y}(0)|$$

有时候也采用互相关系数来刻画两个信号之间的相关性，互相关系数定义为

$$\rho_{xy}(k)=\frac{r_{xy}(k)}{(r_{x}(0)r_{y}(0))^{1/2}}$$

显然，$|\rho_{xy}(k)|\leqslant 1$，$\rho_{xy}(k)$是一个归一化的互相关函数。

如果两个随机信号(或同一随机信号在不同时刻的取值)满足

$$r_{xy}(n_1,n_2)=E[x(n_1)y^{*}(n_2)]=E[x(n_1)]E[y^{*}(n_2)]$$

称 $x(n_1)$和 $y(n_2)$(或同一随机信号在不同时刻的取值)是不相关的，如果对所有n_1和n_2，$x(n_1)$和 $y(n_2)$都是不相关的，则称这两个随机信号是不相关的。如果满足

$$r_{xy}(n_1,n_2)=E[x(n_1)y^{*}(n_2)]=0$$

则称 $x(n_1)$和 $y(n_2)$是正交的。注意到，两个随机变量的正交定义是由他们的互相关定义的。对两个均值为零的随机信号，如果他们是不相关的，必然是正交的。

对于随机信号，用相关性定义正交性是容易理解的，考虑零均值随机信号 $y(n)$，如果我们希望用 y 的当前时刻的值和过去时刻的值及一些 $x(n)$的值预测 y 下一个时刻 $y(n+1)$的值，如果 $x(n_1)$与 $y(n+1)$是不相关的，也就是正交的，我们不期望能从 $x(n_1)$中得到对预测 $y(n+1)$有用的信息，也就是 $y(n+1)$的预测式中 $x(n_1)$项的系数为零。这些讨论也可以用矢量空间的概念来理解。由 y 的过去值和 $x(n)$的一些值构成一个矢量空间 $\mathbf{V}$，其中 $x(n_1)$表示一个矢量空间的一个坐标轴，设 $y(n+1)$是更高维矢量空间的一个矢量，预测 $y(n+1)$实际上是求 $y(n+1)$在矢量空间 $\mathbf{V}$ 的投影，如果 $y(n+1)$与 $x(n_1)$是正交的，则它在 $x(n_1)$方向上的投影必然为零，因此 $y(n+1)$的预测式中不包括 $x(n_1)$项。

例 1.1.3 设 $z(n)=x(n)+y(n)$，$x(n)$，$y(n)$的均值为零，求 $z(n)$的自相关：(1)讨论非平稳的一般情况，(2)讨论 $x(n)$，$y(n)$不相关的情况，(3)讨论平稳情况。

解

(1)
$$\begin{aligned}r_z(n_1,n_2)&=E[(x(n_1)+y(n_1))(x(n_2)+y(n_2))^{*}]\\&=E[x(n_1)x^{*}(n_2)]+E[x(n_1)y^{*}(n_2)]+E[y(n_1)x^{*}(n_2)]+\\&\quad E[y(n_1)y^{*}(n_2)]\\&=r_x(n_1,n_2)+r_{xy}(n_1,n_2)+r_{yx}(n_1,n_2)+r_y(n_1,n_2)\end{aligned}$$

(2) 当 $x(n)$，$y(n)$不相关，且 $x(n)$，$y(n)$的均值为零时，上式简化为

$$r_z(n_1,n_2)=r_x(n_1,n_2)+r_y(n_1,n_2)$$

(3) 在平稳且 $x(n)$，$y(n)$不相关时，$z(n)$的自相关序列进一步简化为

$$r_z(k)=r_x(k)+r_y(k)$$