学者书屋系列

信号分析与数据统计学习

张向君◎编著

经典

哈尔滨工程大

Harbin Engineering

学者书屋系列

信号分析与数据统计学习

张向君　编著

哈尔滨工程大学出版社

内 容 简 介

本书介绍了信号分析与数据统计学习的基本理论、方法和研究进展，并从应用角度，在方法及技术原理等方面进行了详细讨论，对于解决工程实践中的具体问题具有重要指导作用。

本书可作为工程技术人员和高等院校有关专业人员参考用书。

图书在版编目(CIP)数据

信号分析与数据统计学习/张向君编著.—哈尔滨：哈尔滨工程大学出版社，2009.1

ISBN 978-7-81133-394-7

Ⅰ.信…　Ⅱ.张…　Ⅲ.①信号分析-基本知识②统计信号-统计数据-基本知识　Ⅳ.TN911

中国版本图书馆 CIP 数据核字(2009)第 005128 号

出版发行　哈尔滨工程大学出版社
社　　址　哈尔滨市南岗区东大直街 124 号
邮政编码　150001
发行电话　0451-82519328
传　　真　0451-82519699
经　　销　新华书店
印　　刷　哈尔滨工业大学印刷厂
开　　本　787mm×960mm　1/16
印　　张　12.5
字　　数　210 千字
版　　次　2009 年 2 月第 1 版
印　　次　2009 年 2 月第 1 次印刷
定　　价　25.00 元
http://press.hrbeu.edu.cn
E-mail:heupress@hrbeu.edu.cn

前　言

人类已进入信息社会与知识经济时代,信息科学技术迅速发展并广泛应用于各领域。国内外非常重视信号分析与统计学习理论,并对其进行深层次探索,在工程实践中不断提出新理论、新方法。其中,现代信号分析的发展在非线性、自适应两方面结合得越来越紧密,为实现智能化信息处理创造了良好的应用前景;统计学习理论是针对小样本情况发展起来的比较完善的理论体系,它克服了基于渐进理论的传统统计学方法的局限,对于解决工程实践中的具体问题具有重要指导作用。

信号分析与统计学习理论发展到现在,内容非常丰富,本书只就一些重要的基本理论方法进行研究。全书分为两部分:第一部分是信号分析的基础理论和方法,包括第一章连续时间信号及其时域分析,第二章连续时间系统的频域分析,第三章离散时间信号与系统,第四章高分辨时频分析;第二部分是数据统计学习的基础理论和方法,包括第五章统计学习方法,第六章神经网络学习方法,第七章支持向量机学习方法,第八章集成学习。

本书除了反映信号分析和数据统计学习的一些重要基础理论与学术前沿,还包含了作者的研究成果。希望本书对读者有所裨益,对信号分析方法研究和数据统计学习方法研究有所启发。

本书在编写过程中得到了西安交通大学高静怀教授、大庆石油学院唐世伟教授的热心指导,在此表示诚挚的谢意。

信号分析的理论方法与统计学习的理论方法发展得非常迅速,由于作者的学术水平有限,书中难免出现错误和不足,恳请读者批评指正。

作者

2008 年 11 月

目　　录

试读结束，需要全本PDF请购买 www.ertongbook.com

目　录

第1章　连续时间信号与系统

信号是各类信息的运载工具，是某种变化的物理量。不同的信号都包含有一定的意义，这些意义统称为信息，信息中实质性的内容可用信息量度量。在自然、物理、社会等诸多领域中，系统的概念与方法被广泛应用。系统泛指由若干相互作用、相互关联的事物组合而成的具有特定功能的整体。通信、控制系统是信息科学与技术领域的重要组成部分，它们还可以组合成更复杂的系统。描述信号与系统有时域、频域、复频域三种方法。

1.1　信号及其运算

1.1.1　信号的分类

1.确定信号与随机信号

任意由确定时间函数描述的信号，称为确定信号或规则信号。对于这种信号，给定某一时刻后，就能确定一个相应的信号值。如果信号是时间的随机函数，事先将无法预知它的变化规律，这种信号称为不确定信号或随机信号。

2.连续信号与离散信号

一个信号，如果在某个时间区间内除有限个间断点外都有定义，就称该信号在此区间内为连续时间信号，简称连续信号。这里“连续”一词是指在定义域内(除有限个间断点外)信号变量是连续可变的。至于信号的取值，在值域内可以是连续的，也可以是跳变的。

图1-1(a)是正弦信号，其表达式为$f_1(t)=A\sin(\pi t)$，式中，A是常数，其自变量t在定义域$(-\infty,+\infty)$内连续变化，信号在值域$[-A,A]$上连续取值。为了简便起见，若信号表达式中的定义域为$(-\infty,+\infty)$时，则可省去不写。也就是说，凡没有标明时间区间时，均默认其定义域为$(-\infty,+\infty)$。

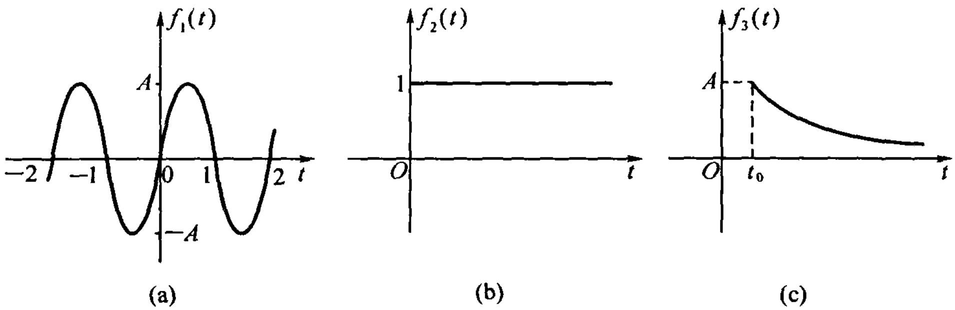

图 1-1 连续信号

图 1-1(b)是单位阶跃信号,通常记为 $\varepsilon(t)$,其表达式为

$$f_2(t) = \varepsilon(t) = \begin{cases} 1 & (t > 0) \\ 0 & (t < 0) \end{cases} \tag{1-1}$$

图 1-1(c)表示一个延时的单边指数信号,其表达式为

$$f_3(t) = \begin{cases} Ae^{-\alpha(t-t_0)} & (t > 0) \\ 0 & (t < 0) \end{cases} \tag{1-2}$$

式中,A 是常数,$\alpha > 0$,信号变量 t 在定义域$(-\infty, +\infty)$内连续变化,信号 $f_3(t)$在值域$[0, A)$上连续取值。注意,$f_3(t)$在 $t = t_0$ 处有间断点。

对于间断点处的信号值一般不作定义,因为这样不会影响分析结果。如有必要,也可按高等数学规定,定义信号 $f(t)$在间断点 t_0 处的信号值等于其左极限 $f(t_{0^-})$与右极限 $f(t_{0^+})$的算术平均值,即

$$f(t_0) = \frac{1}{2}[f(t_{0^-}) + f(t_{0^+})]$$

$$f(t_{0^-}) = \lim_{\xi \to 0} f(t_0 - \xi)$$

$$f(t_{0^+}) = \lim_{\xi \to 0} f(t_0 - \xi)$$

这样,图 1-1 中的信号 $f_2(t)$和 $f_3(t)$也可表示为

$$f_2(t) = \varepsilon(t) = \begin{cases} 1 & t > 0 \\ \dfrac{1}{2} & t = 0 \\ 0 & t < 0 \end{cases} \tag{1-3}$$

$$f_3(t)=\begin{cases}Ae^{-\alpha(t-t_0)} & t>t_0\\ \dfrac{A}{2} & t=t_0\\ 0 & t<t_0\end{cases} \tag{1-4}$$

仅在离散时刻点上有定义的信号称为离散时间信号,简称离散信号。这里“离散”一词表示自变量只取离散的数值,相邻离散时刻点的间隔可以是相等的,也可以是不相等的。在这些离散时刻点以外,信号无定义。信号的值域可以是连续的,也可以是不连续的。

定义在等间隔离散时刻点上的离散信号也称为序列,通常记为 $f(k)$,其中 k 称为序号。与序号 m 相应的序列值 $f(m)$ 称为信号的第 m 个样值。序列 $f(k)$ 的数学表达式可以写成闭式,也可以直接列出序列值或者写成序列值的集合。例如,图 1-2(a)所示的正弦序列可表示为

$$f_1(k)=A\sin\left(\frac{\pi}{4}k\right)$$

随 k 的变化,序列值在值域 $[-A,A]$ 上连续取值。对于图 1-2(b)所示的序列则可表示为

$$f_2(k)=\begin{cases}2 & k=-1,0\\ 1 & k=1\\ -1 & k=2\\ 0 & \text{其他 } k\end{cases}$$

或者

$$f_2(k)=\{\cdots,0,2,\underset{\substack{\uparrow\\ k=0}}{2},1,-1,0,\cdots\}$$

式中,箭头指明 $k=0$ 的位置。同理,图 1-2(c)信号可表示为

$$f_3(k)=\{\cdots,0,A,\underset{\substack{\uparrow\\ k=0}}{A},A,A,0,\cdots\} \tag{1-5}$$

在工程应用中,常常把幅值可连续取值的连续信号称为模拟信号[如图 1-1(a)所示],把幅值可连续取值的离散信号称为采样信号[如图 1-2(a)所示],而把幅值只能取某些规定数值的离散信号称为数字信号[如图 1-2(c)所示]。

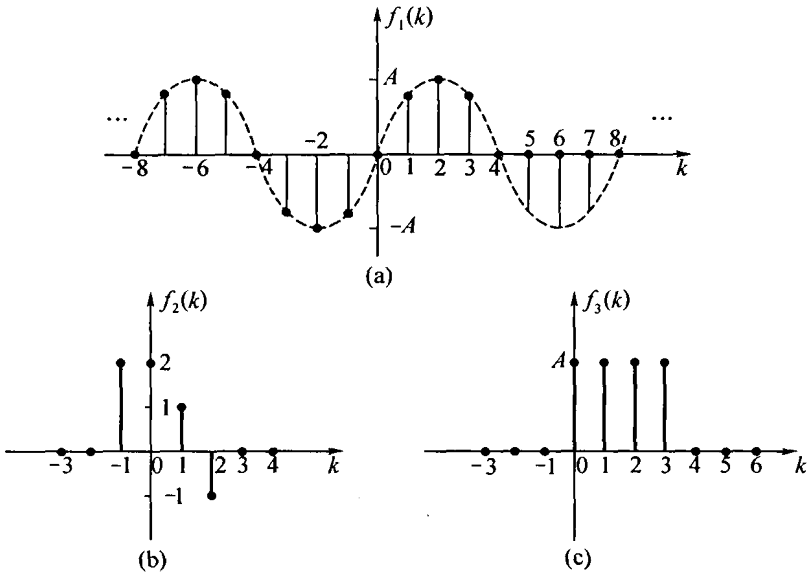

图 1-2 离散信号

为方便起见，有时将信号 $f(t)$ 或 $f(k)$ 的自变量省略，简记为 $f(\cdot)$，表示信号变量允许取连续变量或者离散变量，即用 $f(\cdot)$ 统一表示连续信号和离散信号。

3.周期信号与非周期信号

一个连续信号 $f(t)$，若对所有 t 均有 $f(t)=f(t+mT)$，其中，$m=0,\pm1,\pm2,\cdots$，则称 $f(t)$ 为连续周期信号，满足上式的最小 T 值称为 $f(t)$ 的周期。一个离散信号 $f(k)$，若对所有 k 均有 $f(k)=f(k+mN)$，其中，$m=0,\pm1,\pm2,\cdots$，就称 $f(k)$ 为离散周期信号或周期序列。满足上式的最小 N 值称为 $f(k)$ 的周期。

另外，如果两个周期信号 $x(t)$ 和 $y(t)$ 的周期具有公倍数，则它们的和信号 $f(t)$[即 $f(t)=x(t)+y(t)$]仍然是一个周期信号，其周期是 $x(t)$ 和 $y(t)$ 周期的最小公倍数。

4.能量信号与功率信号

若将信号 $f(t)$ 设为电压或电流，则加载在单位电阻上产生的瞬时功率为 $|f(t)|^2$，在一定的时间区间 $\left[-\frac{\tau}{2},\frac{\tau}{2}\right]$ 内会消耗一定的能量。把该能量对时间区

间取平均,即得信号在此区间内的平均功率。现在将时间区间无限扩展,定义信号 $f(t)$ 的能量 E 为

$$E = \lim_{\tau \to \infty} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} |f(t)|^2 \mathrm{d}t \quad P = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} |f(t)|^2 \mathrm{d}t \tag{1-6}$$

如果在无限大时间区间内,信号的能量为有限值(此时平均功率 $P=0$),就称该信号为能量有限信号,简称能量信号。如果在无限大时间区间内,信号的平均功率为有限值(此时信号能量 $E=\infty$),则称此信号为功率有限信号,简称功率信号离散信号 $f(k)$,$f(k)$ 的能量定义为

$$E = \sum_{k=-\infty}^{+\infty} |f(k)|^2 \tag{1-7}$$

1.1.2 单位斜坡信号

单位斜坡信号的表达式为

$$r_1(t) = \begin{cases} 0 & \left(t < -\dfrac{\tau}{2}\right) \\ \dfrac{1}{\tau}t + \dfrac{1}{2} & \left(-\dfrac{\tau}{2} \leqslant t \leqslant \dfrac{\tau}{2}\right) \\ 1 & (t > 0) \end{cases} \tag{1-8}$$

单位斜坡信号的波形如图 1-3 所示。

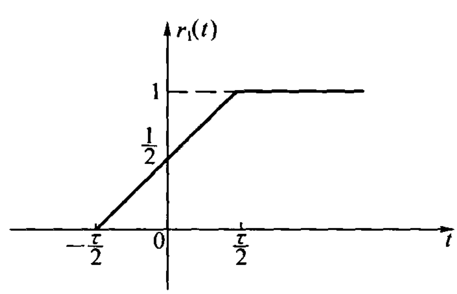

图 1-3 单位斜坡信号

试读结束,需要全本PDF请购买 www.ertongbook.com

1.1.3 单位阶跃信号

单位阶跃信号的定义为

$$\varepsilon(t)=\begin{cases}0 & (t<0)\\ 1 & (t>0)\end{cases} \tag{1-9}$$

单位阶跃信号的波形如图 1-4 所示。在 $t=0$ 处的函数值未作定义。

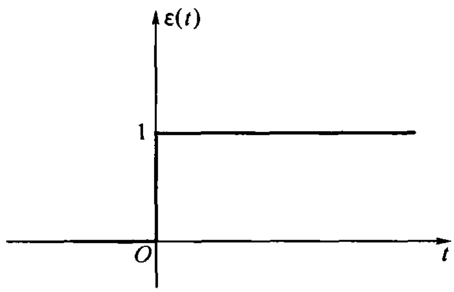

图 1-4 单位阶跃信号

需要提醒大家注意的是:有些书用 $U(t)$或 $u(t)$表示单位阶跃信号,但我们知道 $U(t)$和 $u(t)$更多的时候是用来表示电压的,因此,为了与电压表示符号相区别,我们用 $\varepsilon(t)$表示单位阶跃信号。

一般规定,若函数在 $t=t_0$ 时刻出现跳变,则可以用该点左右极限的平均值作为该点的函数值。那么,$\varepsilon(t)$在 $t=0$ 时刻的值可以表示为

$$\varepsilon(0)=\frac{\varepsilon(0^-)+\varepsilon(0^+)}{2}=\frac{1}{2} \tag{1-10}$$

这样,单位阶跃信号可以用符号信号和直流信号表示为

$$\varepsilon(t)=\frac{1}{2}+\frac{1}{2}\operatorname{sgn} t \tag{1-11}$$

单位阶跃信号还可看作是单位斜坡信号的极限情况。从图 1-3 可知,若单位斜坡信号的斜坡时间区间为 τ,则当斜坡时间区间趋于零时,即 $\tau\to 0$ 时,单位斜坡信号就变成了单位阶跃信号。

单位阶跃信号的一个重要用途是用来表示一些其他信号,以简化信号的运算

及其表达式。

1.1.4　单位冲激信号

狄拉克(Dirac)给出单位冲激信号的一种定义方式为

$$\begin{cases}\delta(t) = 0 \qquad (t \neq 0) \\ \int_{-\infty}^{+\infty} \delta(t)\mathrm{d}t = 1\end{cases} \tag{1-12}$$

单位冲激信号函数值在原点以外处处为零,而$\int_{-\infty}^{+\infty} \delta(t) = 1$表示该信号波形下的面积为1。单位冲激信号$\delta(t)$可以看作为在较短的时间内产生很大能量的物理现象的理想化模型。比如自然界中的电闪雷击、地震,工业生产中的强电火花,生活中用榔头敲钉子等现象都可看作为现实生活中的冲激信号。

1.抽样性质

若函数$f(t)$在$t=0$处连续,且处处有界,则有

$$f(t)\delta(t) = f(0)\delta(t) \tag{1-13}$$

上式说明,$f(t)$与$\delta(t)$相乘的结果是强度为$f(0)$的冲激函数。对上式两边同时作积分,有

$$\int_{-\infty}^{+\infty} f(t)\delta(t)\mathrm{d}t = \int_{-\infty}^{+\infty} f(0)\delta(t)\mathrm{d}t = f(0)\int_{-\infty}^{+\infty} \delta(t)\mathrm{d}t = f(0) \tag{1-14}$$

将上式推广为

$$\int_{-\infty}^{+\infty} f(t)\delta(t-t_0)\mathrm{d}t = \int_{-\infty}^{+\infty} f(t_0)\delta(t-t_0)\mathrm{d}t = f(t_0) \tag{1-15}$$

式(1-14)和式(1-15)说明了冲激函数的抽样性质,即一个连续有界的函数$f(t)$与单位冲激函数$\delta(t)$相乘并从$(-\infty,+\infty)$作积分,可以得到$f(t)$在$t=0$时刻的函数值$f(0)$,筛选出函数值$f(0)$。若$\delta(t)$延迟t_0个单位,则筛选出$f(t_0)$。

2.偶函数性质

$\delta(t)$是偶函数,即$\delta(-t)=\delta(t)$。

证明:考虑积分$\int_{-\infty}^{+\infty} f(t)\delta(-t)\mathrm{d}t$,并令$t=-\tau$,则有

$$\int_{-\infty}^{+\infty} f(t)\delta(-t)\mathrm{d}t = -\int_{+\infty}^{-\infty} f(-\tau)\delta(\tau)\mathrm{d}t = \int_{-\infty}^{+\infty} f(-\tau)\delta(\tau)\mathrm{d}t = f(0) \tag{1-16}$$

结合 $\int_{-\infty}^{+\infty} f(t)\delta(t)\mathrm{d}t = f(0)$，从而有

$$\delta(-t) = \delta(t)$$

3.单位冲激函数的导数

单位冲激函数的一阶导数记作 $\delta'(t)$，即 $\delta'(t) = \dfrac{\mathrm{d}\delta(t)}{\mathrm{d}t}$，我们称之为单位冲激偶信号，其波形如图 1-5 所示。

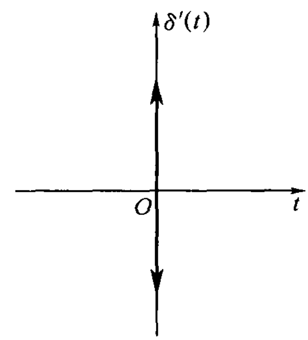

图 1-5 单位冲激偶信号

冲激偶信号的一个重要性质是

$$\int_{-\infty}^{+\infty} \delta'(t)f(t)\mathrm{d}t = -f'(0) \qquad (1-17)$$

其中，$f(t)$在 $t=0$ 处连续，$f'(0)$为 $f(t)$的一阶导函数在 $t=0$ 处的取值。

证明：

$$\int_{-\infty}^{+\infty} \delta'(t)f(t)\mathrm{d}t = \int_{-\infty}^{+\infty} f(t)\mathrm{d}\delta(t) = f(t)\delta(t)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{+\infty} \delta(t)f'(t)\mathrm{d}t$$
$$= -f'(0)\int_{-\infty}^{+\infty} \delta(t)\mathrm{d}t = -f'(0)$$

类似地还有

$$\int_{-\infty}^{+\infty} \delta^{(n)}(t)f(t)\mathrm{d}t = -(-1)^n f^{(n)}(0)$$

冲激偶信号的另一个重要性质是

$$\int_{-\infty}^{+\infty} \delta'(t)\mathrm{d}t = 0$$

4.单位冲激信号的积分等于单位阶跃信号

由狄拉克定义式可知

$$\int_{-\infty}^{t} \delta(\tau)\mathrm{d}\tau = \begin{cases} 0 & (t<0) \\ 1 & (t>0) \end{cases} \qquad (1-18)$$

对照单位阶跃信号的定义式，可得出

$$\int_{-\infty}^{t} \delta(\tau)\mathrm{d}\tau = \varepsilon(t) \qquad (1-19)$$

与之相对应，单位阶跃信号的导数等于单位冲激信号

$$\delta(t) = \frac{\mathrm{d}\varepsilon(t)}{\mathrm{d}t} \qquad (1-20)$$

1.1.5　信号的运算

1.加法运算

任一瞬间的和信号值 $y(t)$或 $y(k)$等于同一瞬间相加信号瞬时值的和,即

$$y(t)=f_1(t)+f_2(t)$$

或

$$y(k) = f_1(k) + f_2(k) \tag{1-21}$$

2.乘法运算

任一瞬间的乘积信号值 $y(t)$或 $y(k)$等于同一瞬间相乘信号瞬时值的积,即

$$y(t) = f_1(t)f_2(t)$$
$$y(k) = f_1(k)f_2(k) \tag{1-22}$$

3.数乘(标乘)

信号 $f_1(t)$或 $f_1(k)$和一个常数 a 相乘的积,即

$$y(t) = af_1(t)$$
$$y(k) = af_1(k) \tag{1-23}$$

4.微分

信号的微分是指信号对时间的导数,可表示为

$$y(t) = \frac{\mathrm{d}f(t)}{\mathrm{d}t} = f'(t) \tag{1-24}$$

5.积分

信号的积分是指信号在区间$(-\infty,t)$上的积分,可表示为

$$y(t) = \int_{-\infty}^{t} f(\tau)\mathrm{d}\tau = f^{(-1)}(t) \tag{1-25}$$

图 1-6 是信号积分的一个例子。

6.反转

以变量 $-t$ 代替$f(t)$中的独立自变量 t,可得反转信号 $f(-t)$。它是$f(t)$以纵轴($t=0$)为转轴作 180°反转而得到的信号波形,如图 1-7 所示。

7.平移

以变量$(t-t_0)$代替信号 $f(t)$中的独立变量 t,得信号 $f(t-t_0)$,它是信号 $f(t)$沿时间轴平移 t_0 的波形。这里 $f(t)$与 $f(t-t_0)$的波形形状完全一样,只是在位置

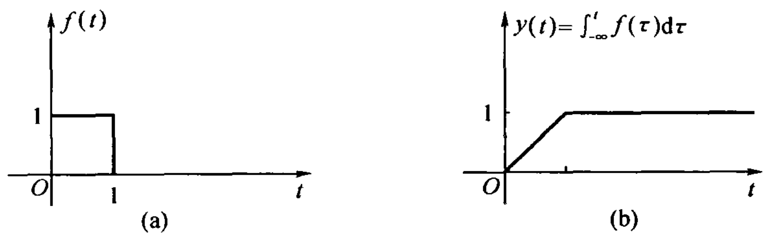

图 1-6 信号的积分

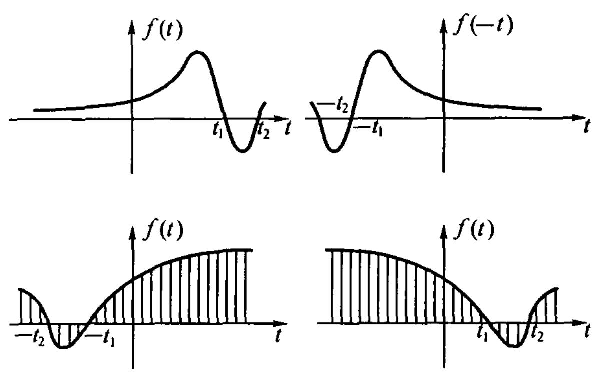

图 1-7 离散时间信号及反转波形

上移动了 t_0(t_0 为一实常数)。$t_0>0$,$f(t)$右移;$t_0<0$,$f(t)$左移;平移距离为$|t_0|$。

图 1-8 表示连续时间信号的平移。这类信号在雷达、声呐和地震信号处理中经常遇到。利用位移信号 $f(t-t_0)$和原信号 $f(t)$在时间上的迟延,可以探测目标和震源的距离。

8.展缩(尺度变换)

以变量 at 代替$f(t)$中的独立变量 t 可得$f(at)$,它是 $f(t)$沿时间轴展缩(尺度变换)而成的一个新的信号函数或波形。信号 $f(at)$中,a 为常数,$|a|>1$ 时表示 $f(t)$沿时间轴压缩成原来的$1/|a|$倍;$|a|<1$ 时表示 $f(t)$沿时间轴扩展为原来的$1/|a|$倍。

例如,图 1-9(a),(b),(c)分别表示 $f(t)$,$f(2t)$及 $f(t/2)$的波形。

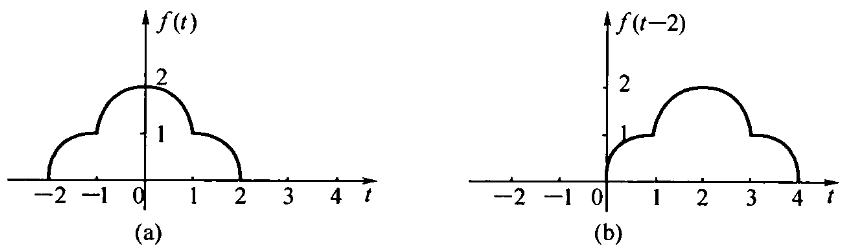

图 1－8　连续时间信号的平移

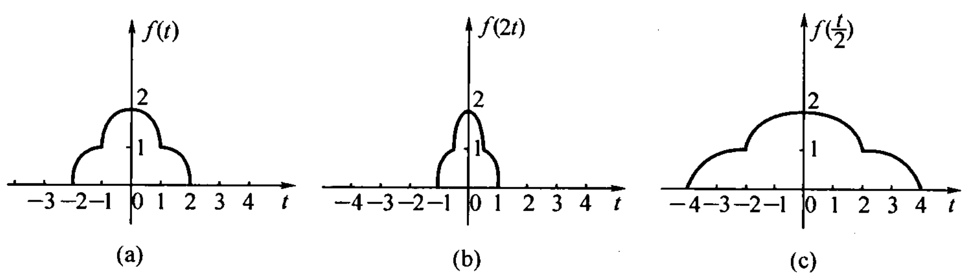

图 1－9　$f(t)$、$f(2t)$ 及 $f(t/2)$ 的波形

9.综合变换

以变量$(at+b)$代替 $f(t)$中的独立变量 t，可得一新的信号函数 $f(at+b)$。当 $a>0$时，$f(at+b)$是 $f(t)$沿时间轴展缩、平移后的信号波形；当 $a<0$ 时，$f(at+b)$是 $f(t)$沿时间轴展缩、平移和反转后的信号波形。

1.2　连续时间系统的时域分析

1.2.1　微分方程的经典解法

一般 n 阶 LTI 系统的微分方程为

$$\frac{d^n y(t)}{dt^n}+a_1\frac{d^{n-1}y(t)}{dt^{n-1}}+\cdots+a_{n-1}\frac{dy(t)}{dt}+a_n y(t)=$$

试读结束，需要全本PDF请购买 www.ertongbook.com