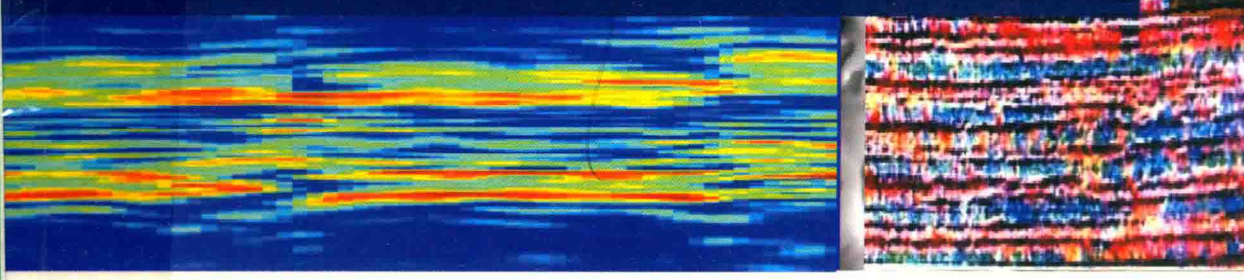
地球物理信息处理基础

Fundamentals of Geophysical Information Processing

刘喜武　张远银　孟凡顺　◎著



石油工业出版社

国家自然科学基金委员会—中国石油化工股份有限公司石油化工联合
基金项目“页岩油气层及裂缝多尺度多物理正反演(U1663207)”资助

地球物理信息处理基础

刘喜武　张远银　孟凡顺　著

石油工业出版社

内容提要

本书以地球物理信号处理和数据分析中应用的现代信号处理和统计估计为基础，结合它们在地球物理数据处理和分析中的应用，介绍它们的基本原理和实现方法，但不追求繁杂的数学论证，尽量做到深入浅出、通俗易懂。具体包括随机信号基础，Fourier 分析，采样、Z 变换和褶积，离散 Fourier 变换，Fourier 变换有关的变换，地球物理数字滤波和地震反射序列反褶积等七个章节。

本书可供从事地球物理勘探研究工作的科研人员参考使用。

图书在版编目(CIP)数据

地球物理信息处理基础/刘喜武，张远银，孟凡顺著.—北京：石油工业出版社，2017.10

ISBN 978-7-5183-2109-4

Ⅰ.①地… Ⅱ.①刘… ②张… ③孟… Ⅲ.①地球物理学-信息处理 Ⅳ.①P3-39

中国版本图书馆 CIP 数据核字(2017)第 228755 号

出版发行：石油工业出版社
(北京安定门外安华里 2 区 1 号楼 100011)
网 址：www.petropub.com
编辑部：(010)64523708 图书营销中心：(010)64523633
经 销：全国新华书店
印 刷：北京晨旭印刷厂

2017 年 10 月第 1 版 2017 年 10 月第 1 次印刷
787×1092 毫米 开本：1/16 印张：13
字数：230 千字

定价：100.00 元
(如出现印装质量问题，我社图书营销中心负责调换)

版权所有，翻印必究

前　　言

信息与信号处理是信息科学近十几年来发展最快的学科之一。有些学者将其视为一个国家科学水平发展的标志性学科之一，从某种意义上说，这也许并不为过。信息与信号处理在通信、雷达、声呐、地质勘探、医学仪器、振动工程、射电天文、控制、电子工程等领域十分重要，并得到广泛应用。地球物理信息处理，就是利用计算机分析和处理地球物理信号和数据，信号处理和数据分析的理论和方法是其重要的基础。

本书以地球物理信号处理和数据分析中应用的现代信号处理和统计估计为基础，结合它们在地球物理数据处理和分析中的应用，介绍其基本原理和实现方法。本书内容不追求繁杂的数学论证，尽量做到深入浅出、通俗易懂。

全书共八章。第一章为绪论。第二章随机信号基础，介绍现代信号处理中的常用内容和信号处理的一些数学基础。第三章 Fourier 分析，介绍 Fourier 分析的一些基本原理，如 Fourier 级数、连续 Fourier 变换等。第四章采样、Z 变换和褶积，介绍处理离散时间序列和线性系统重要的方法——Z 变换和褶积，第三章介绍的 Fourier 变换处理连续信号，而 Z 变换处理离散信号的采样，和后面介绍的离散 Fourier 变换关系密切。第五章离散 Fourier 变换，介绍从 Z 变换到离散 Fourier 变换的转化，离散 Fourier 变换的计算和实现。第六章 Fourier 变换有关的变换，简要介绍 Laplace 变换，以及与 Fourier 变换直接相关且有重要应用价值和物理意义的 Hilbert 变换、Radon 变换；给出小波变换和分数阶 Fourier 变换基本概念。第七章地球物理数字滤波，介绍地球物理中应用的数字滤波方法。第八章地震反射序列反褶积，又称地震道反演，介绍地震勘探反射系数反褶积的计算和数值实现。

本书可供高等学校地球探测与信息技术和固体地球物理学专业研究生，地球信息科学与技术、勘查技术与工程和固体地球物理学等专业高年级本科生，以及信息与计算科学、通信、自动化等专业学生学习参考，也可作为从事信号处理和地球物理勘探的工程技术人员学习地震信号处理的基础材料。

限于笔者水平和经验，书中难免存在不足，敬请读者批评指正。

目　　录

此为试读，需要完整PDF请访问：www.ertongbook.com

第一章　绪　　论

确定性信号处理的理论和方法，讨论连续和离散确定性信号的时域、频域表示方法，以及系统对信号所做的各种处理是信号处理（数据分析）的基础，包括离散采样和重构、信号与系统、线性时不变系统、Fourier 变换、Z 变换、数字滤波器等主要内容，研究的对象是确定性信号和线性、时不变、因果（物理可实现）、最小相位系统。而现代信号处理则主要研究随机（统计）信号和非线性、非因果、非最小相位系统。对随机信号进行系统和科学的分析及处理，显然具有更为重要的意义，因为确定性信号实际上只是随机信号的某种特例。传统的统计信号处理有 3 个基本假设：线性、高斯性和平稳性，主要内容包括离散随机信号基础、估计理论、最优滤波、最小二乘法滤波、功率谱估计、自适应滤波。现代统计信号处理则以非线性、非高斯和非平稳信号作为分析和处理对象，特别是非平稳信号处理与自适应信号处理两者结合越来越紧密，主要内容包括高阶谱分析、高阶累积量、分数低阶矩或分数低阶统计量、时频分析方法（特别是自适应的时频分析与分数阶广义时频分析）、小波变换、非平稳随机信号谱估计、独立分量分析等。这些内容反映了当今信号处理国际学术发展的前沿。

信号处理在数学上就是一种数学变换和反演估计方法。信号处理一定涉及数学变换（如 Fourier 变换、Z 变换、Laplace 变换、短时 Fourier 变换、小波变换、希尔伯特变换、各种时频变换方法、分数阶 Fourier 变换、KL 变换、Radon 变换等）和反演方法（最小方差估计、最小二乘反演、最大似然估计、各种滤波和反演算法等）。信号处理致力于信号和数据的有效表示。从数学的观点来看，通过在函数的完备集中展开信号，就可以实现信号的不同表示。一种特别表示的重要意义是可以更好地理解信号的特征，这种表示是由实际上或对眼前情况重要的物理量表示其特征。

地球物理信息处理就是利用计算机分析地球物理资料。信号处理的理论和方法是地球物理信息处理的重要基础，特别是现代信号处理理论与方法应用于各种地球物理信息和数据处理问题。地球物理数据处理是从研究时间域采样的滤波理论和谱分析开始的。而随机信号处理的一个重要问题是波形估计，也称滤波问题。滤波一词和含义远不限于"选频"这一狭窄的概念。凡是对时间序列（随机信号的一次实现）进行的处理以改善对信号的理解过程均称为滤波。平稳随机信号最优估计归结为维纳（Winer）滤波，即地球物理数据处理中的最小平方滤波，按照均方误差最小准则设计滤波器，例如地震勘探最小平方反褶积、预测反褶积、同态反褶积、最小熵与最大熵反褶积。非平稳随机信号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学系统的状态变量进行估计和预测，地震勘探最大似然反褶积就是以 Kalman 滤波为基础的方法，递推结构使得 Kalman 滤波可以方便地跟踪非平稳的状态变化，并可扩展到非线性系统。Winer 滤波给出了线性滤波在统计意义下的最优解，但 Winer 滤波的实现需要对输入信号进行先验统计，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预置，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进行调节，估计结果将收敛到逼近于最优滤波器（Winer 滤波器）的性能。这类线性滤波器称为自适应滤波器，如地震勘探中的盲反褶积和盲分离方法。

在线性自适应滤波和功率谱估计中仅使用随机信号的二阶矩。二阶矩对高斯过程能完全表示，但对非高斯过程二阶矩的表示是不完整的。那么，就可以利用高阶矩（离散统计）获得随机信号更多有用的信息，进行非高斯随机信号处理。例如，在地震子波估计中，利用二阶矩和功率谱进行子波估计是相位盲的，即不包含相位信息，只能处理最小相位子波，而利用高阶矩和高阶谱分析可以提取子波和相位信息，实现非最小相位子波估计（系统辨识）和盲反褶积。最近兴起的盲信号处理也是利用高阶统计学方法，如独立分量分析（ICA）、盲源分离（BSS）和盲信号提取（BSE）等。

此外，地球物理信号在地层中传播，由于吸收衰减，实际是时变的，即非平稳随机过程。要想获得局域瞬变过程特性得到精细分析结果，必须利用具有时频局部化特性的分析工具——时频分析（Time - Frequency Analysis），这也是其他处理分析方法的基础。傅里叶变换（Fourier Transform）是线性系统分析最重要的工具，必须在整个时间轴上对 $f(t)$ 和 $e^{i\omega t}$ 进行混合。因此，傅里叶变换无法抽取信号的局域性质，即没有能力抽取或定位信号在某个时间附近的瞬变特性。为此，人们研究各种时频分析方法，进行局域时间和频率联合分析，如短时 Fourier 变换及离散 Gabor 展开、Wigner 分布、小波变换、Cohen 类时频分布、自适应核函数时频分析、基于分数阶 Fourier 变换的广义时频分析，以及局域波分解方法（Hilbert - Huang 变换、经验模式分解）等方法。它们广泛用于地球物理信号分析和处理中（如地震波能量补偿、时变滤波去噪、精确检测地震纵波到达时），用于检测地层变化、提高地震勘探资料对薄层的识别能力及研究地震层序和旋回分析等方面。

综上，信号分析和数据处理的基础理论和方法作为地球物理数据处理（分析）的基础，主要用于地球物理数据的信息改进，如数字滤波、反褶积、初静校正、叠加、成像、去噪、提高分辨率等，近年来大力发展了信息提取、分析和预测方法，如各种地震属性分析、频谱分解、地震相分析、速度分析、储层参数分析及预测等。

信号处理中涉及的统计数据分析、时间序列分析和反演算法等具有更广泛和深刻的内容，实际应用中的关键和核心可归结为科学计算问题。

现代信号处理方法在地球物理中的应用非常广泛。本书主要内容为以现代信号处理和反演算法为基础的现代地球物理数据处理的基本原理和方法，要求的线性代数、信号与系统、矩阵计算、概率论与数理统计、随机过程和时间序列分析、现代信号处理统计估计等基础内容具有自包容特点。同时，书中部分章节的方法均提供科 Matlab 程序，便于读者更好地理解和掌握。

第二章　随机信号基础

第一节　数学预备知识

一、离散序列与 n 维(空间)向量

1. n 维空间

由 n 个有顺序的实数构成的实数组 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 的全体构成的集合,称为 n 维空间,实数组 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 称为 n 维空间的一点,或一个离散序列,$x_1,x_2,\cdots,x_n$ 称为这个实数组的分量或坐标;由于点 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 又可表示从原点$(0,0,\cdots,0)$出发的一个 n 维向量,所以 n 维空间又称 n 维向量空间。

设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 为 n 个线性无关的 n 维向量,$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 是任意 n 维向量,则 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 一定能表示称 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 的线性组合。

$$\boldsymbol{x}=\sum_{i=1}^{n}x_i\boldsymbol{e}_i \tag{2-1}$$

其中,坐标 $x_1,x_2,\cdots,x_n$ 为表示的系数;$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 为 n 维空间的一组基或基向量。

2. 向量的范数(模)

如果 $\boldsymbol{V}$ 是数域 $\boldsymbol{K}$ 上的线性空间,且对于 $\boldsymbol{V}$ 的任一向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$,对应一个实值的泛函数 $\|\boldsymbol{x}\|$,它满足:

(1)非负性。当 $\boldsymbol{x}\neq 0$ 时,$\|\boldsymbol{x}\|>0$;当 $\boldsymbol{x}=0$ 时,$\|\boldsymbol{x}\|=0$;

(2)齐次性。$\|a\boldsymbol{x}\|=a\|\boldsymbol{x}\|,a\in\boldsymbol{K},\boldsymbol{x}\in\boldsymbol{V}$;

(3)三角不等式。$\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|,\boldsymbol{x},\boldsymbol{y}\in\boldsymbol{V}$;

则称 $\|\boldsymbol{x}\|$ 为 $\boldsymbol{V}$ 上向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 的范数(模)。

最常见的范数有以下 3 种:

(1)2—范数(l_2),$\|\boldsymbol{x}\|_2=\sqrt{|x_1|^2+|x_2|^2+\cdots+|x_n|^2}$

(2)∞—范数(l_∞),$\|\boldsymbol{x}\|_\infty=\max|x_i|$

(3)1—范数(l_1),$\|\boldsymbol{x}\|_1=|x_1|+|x_2|+\cdots+|x_n|$

一般定义 p—范数或 l_p 范数为

$$\|\boldsymbol{x}\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}} \tag{2-2}$$

3. l_2 向量空间

向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 若满足 $\|\boldsymbol{x}\|_2=\sqrt{\sum\limits_{i=1}^{n}|x_i|^2}<\infty$，则称由 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 全体构成的向量空间为 l_2 向量空间。

4. L_2 空间

复值函数 $f(t)$ 若满足能量有限条件 $\int_{-\infty}^{+\infty}|f(t)|^2\mathrm{d}t<\infty$，则称由 $f(t)$ 全体构成的向量空间为 L_2 空间。

l_2 向量空间中的每一个向量是 L_2 空间某一函数的 Fourier 系数所组成的离散序列。

5. 两个向量的内积(点积)与正交性

设两个 n 维向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 和 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$，定义内积为

$$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_iy_i \tag{2-3}$$

显然，$(\boldsymbol{x},\boldsymbol{x})=\|\boldsymbol{x}\|^2$。根据二维向量的正交性，定义 n 维向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 和 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$ 的正交性。

$$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_iy_i=0 \tag{2-4}$$

设 n 维基向量 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$，彼此正交，其中 $\boldsymbol{e}_k=(0,\cdots,0,1_k,0,\cdots,0)$，$1_k$ 表示第 k 个元素是 1，则 $\boldsymbol{x}=\sum\limits_{i=1}^{n}x_i\boldsymbol{e}_i=(x_1,x_2,\cdots,x_n)$，且 $x_i=(\boldsymbol{x},\boldsymbol{e}_i)$，表明坐标或表示的系数等于向量与相应的基向量的内积。

将两个向量序列的正交性推广到函数正交性。两个定义在区间 $[a,b]$ 上的函数 $u_1(t)$ 和 $u_2(t)$ 的正交性定义为

$$\int_a^b u_1(t)u_2(t)\mathrm{d}t=0 \tag{2-5}$$

区间 $[a,b]$ 称为正交区间。

二、线性变换与矩阵

线性空间的维数是构成该线性空间线性无关向量组所含向量的最大个数。线性空间中的元素满足加法和数乘的封闭性。设 n 维线性空间中，向量 $\boldsymbol{y}=(y_1,y_2,\cdots,y_m)(m\leqslant n)$ 的每一个分量都可以用向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 分量线性表示，即有

$$\begin{cases} y_1=a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n \\ y_2=a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n \\ \vdots \\ y_m=a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n \end{cases} \tag{2-6}$$

称为由变量 $x_1,x_2,\cdots,x_n$ 到变量 $y_1,y_2,\cdots,y_m$ 的线性变换，取决于系数构成的矩阵：

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{2-7}$$

于是

$$\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x} \tag{2-8}$$

线性变换和矩阵之间建立了一一对应的关系，所以研究线性变换，只要研究矩阵的结构即可。

若矩阵满足 $\boldsymbol{A}^T=\boldsymbol{A}^{-1}$，则称 $\boldsymbol{A}$ 为正交矩阵。对于线性表示（变换）$\boldsymbol{x}=\sum_{i=1}^{n}x_i\boldsymbol{e}_i=(x_1,x_2,\cdots,x_n)^T$，也可以写成矩阵形式：

$$\boldsymbol{x}=[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n \tag{2-9}$$

即 $\boldsymbol{x}=\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}\boldsymbol{x}$，矩阵 $\boldsymbol{A}=[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n]=\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$作为线性变换，每一行就是基向量 $\boldsymbol{e}_i$，$i=1,2,\cdots,n$，而且 $\boldsymbol{A}$ 为正交矩阵，线性变换也是正交的。基向量是彼此正交的，响应的线性表示（变换）也是正交的。

如果有这样一个 n 维的非零向量 $\boldsymbol{\varphi}$，由 n 阶方阵定义 $\boldsymbol{A}$ 的所定义的线性变换作用后，只压缩或伸长列向量，则这个列向量 $\boldsymbol{\varphi}$ 称为方阵 $\boldsymbol{A}$ 对应于压缩或伸长系数 λ 的特征向量，而压缩或伸长系数 λ 称为方阵 $\boldsymbol{A}$ 的特征值，即

$$\boldsymbol{A}\boldsymbol{\varphi}=\lambda\boldsymbol{\varphi} \tag{2-10}$$

或

$$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{\varphi}=0 \tag{2-11}$$

记矩阵

$$\boldsymbol{E}=(\boldsymbol{\varphi}_1,\cdots,\boldsymbol{\varphi}_n) \tag{2-12}$$

公式(2-12)为本征向量矩阵，是正交矩阵。

记矩阵

此为试读，需要完整PDF请访问：www.ertongbook.com

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \tag{2-13}$$

为本征向量矩阵。则有

$$\boldsymbol{A} = \boldsymbol{EDE}^T \tag{2-14}$$

称为方阵 $\boldsymbol{A}$ 的本征分解(EVD)。

对于一般的矩阵(非方阵),本征分解则推广为奇异值分解(SVD),SVD 在最小二乘反演、广义逆矩阵的计算、多元统计估计问题、最优化问题和特征值问题方面都有十分重要的应用。设矩阵 $\boldsymbol{A}$ 为 $m \times n$ 阶的矩阵,则总存在 $m \times m$ 阶正交矩阵 $\boldsymbol{U}$ 和 $n \times n$ 阶正交矩阵 $\boldsymbol{V}$,使得

$$\boldsymbol{U}^T\boldsymbol{AV} = \boldsymbol{D} \tag{2-15}$$

其中,$\boldsymbol{D} = \begin{pmatrix} \sum & 0 \\ 0 & 0 \end{pmatrix}$,$\sum = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r)$,而 $\sigma_i = \sqrt{\lambda_i} (i = 1,2,\cdots,r)$,$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r > 0$ 是矩阵 $\boldsymbol{AA}^T$ 或 $\boldsymbol{A}^T\boldsymbol{A}$ 的所有非零特征值,$\sigma_i = \sqrt{\lambda_i} (i = 1,2,\cdots,r)$ 称为矩阵 $\boldsymbol{A}$ 的奇异值。而

$$\boldsymbol{A} = \boldsymbol{UD}\,\boldsymbol{V}^T \tag{2-16}$$

称为矩阵 $\boldsymbol{A}$ 的奇异值的奇异值分解。

三、正交变换与正交函数系

1. 正交变换

在线性变换 $\boldsymbol{y} = \boldsymbol{Ax}$ 中,若 $\boldsymbol{A}$ 为正交矩阵,则称这种线性变换为正交变换。一般各种正交变换能在不同程度上减少随机向量 $\boldsymbol{x}$ 的相关性,从而有利于信号处理,这是正交变换在信号处理中的本质作用。

2. 正交函数系

Fourier 级数是对周期确定性信号(连续函数)的一种正交变换,它把信号分解成相互正交的分量的加权和。对随机信号的任意一次具体实现也可以做类似的分解,当然此时的系数也是随机变量。采用不同的确定性正交函数系(基函数)$\{\varphi_i(t), i = 0,1,2,\cdots\}$,就得到不同的正交分解,如 Fourier 变换、小波变换、K—L 变换、沃希变换,哈尔变换等。

对于函数系 $\{\varphi_i(t), i = 0,1,2,\cdots\}$,都属于 L_2 空间,且满足:

$$\int_a^b \varphi_i(t)\varphi_j(t)\,\mathrm{d}t = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases} \quad i,j = 1,2,\cdots,n \tag{2-17}$$

称函数系 $\varphi_i(t), i = 1,2,\cdots$ 在区间 $[a,b]$ 上为归一化正交函数系。

设 $f(t)$ 为定义在区间 $[a,b]$ 上的实值函数(信号),可用函数系(基)展开(表示)为

$$f(t) = \sum_{i=0}^{\infty} c_i\varphi_i(t) \tag{2-18}$$

其中,函数系$\{\varphi_i(t), i=0,1,2,\cdots\}$称为基函数。基函数与基向量相对应。正交的基函数是对函数的变换,是正交线性变换,起到解相关的作用。展开的系数相当于前面讨论基向量时的坐标。根据定义,坐标$x_i = (\boldsymbol{x}, \boldsymbol{e}_i), i = 1,\cdots,n$,很容易求得正交函数展开的系数$c_j = \int_a^b f(t)\varphi_j(t)\mathrm{d}t, j = 1,\cdots,n$。即式(2-18)两边同时乘$\{\varphi_j(t), j=0,1,2,\cdots\}$,在$[a,b]$积分,根据基正交性,可得

$$\int_a^b f(t)\varphi_j(t)\mathrm{d}t = \int_a^b \sum_{i=0}^{\infty} c_i\varphi_i(t)\varphi_j(t)\mathrm{d}t = \sum_{i=0}^{\infty} c_i\int_a^b \varphi_i(t)\varphi_j(t)\mathrm{d}t = c_j, j = 0,1,\cdots,n$$

即有

$$c_j = \int_a^b f(t)\varphi_j(t)\mathrm{d}t, j = 0,1,\cdots,n \tag{2-19}$$

3. 函数系的封闭性

对任意分段连续函数$x(t)$均可用函数系$\{\varphi_n(t), n=0,1,2,\cdots\}$中有限个函数系的线性组合平均逼近到任意程度,当$n$充分大时,$\int_a^b \left[x(t) - \sum_{i=0}^{n} c_i\varphi_i(t)\right]^2 \mathrm{d}t$可以小于任意给定的整数,称函数系是封闭的。显然,函数的封闭性不一定只对正交系才成立。

4. 正交函数系的完备性

当$x(t) \in L_2$空间(平方可积函数)与正交函数系$\{\varphi_n(t), n=0,1,2,\cdots\}$中的每个函数都正交,则必有$x(t) \equiv 0$,称正交函数系$\{\varphi_n(t), n=0,1,2,\cdots\}$是完备的。$L_2$空间的归一正交系、完备性和封闭性是等价的,只要是完备正交函数系,就可以作为L_2空间的基。

第二节 信号与系统的基本关系

一、信号

携带信息的物理过程称为信号,它是传递信息的函数。信号处理在数学上可以看成具有某种特性的数学变换或反演。信号可以表示成一个或几个独立自变量的函数。对自变量和函数的取值不同,以时间序列为例,可是把信号分成以下 4 种形式:

(1)模拟信号$x(t)$,是指连续时间、连续幅值的信号;

(2)数字信号$x(n)$,是指离散时间、离散幅值的信号;

(3)连续时间信号$x(t)$,是指时间连续,而幅值可以连续也可以离散的信号(常与模拟信号混同);

(4)离散时间信号$x(n\Delta t)$,是指时间离散,而幅值连续的信号。如果在信号处理中暂不

考虑时间量化问题，这种信号也可用 $x(n)$ 表示。

上述信号形式中离散时间信号（时间序列）可以通过对时域连续函数取样得到。采样间隔 Δt 一般为等间隔。

二、系统

一般凡是反映因果关系的装置或运算都可称为系统，或称为滤波器。从数学的角度看，系统不过是具有某种性质的算子，输入信号经过系统（算子、滤波器）作用，得到输出信号，即

$$y(t) = \sigma\{x(t)\} \tag{2-20}$$

只有一个输入和一个输出的系统，称为单道滤波器；其他情况称为多道滤波器，包括多道输入单道输出、单道输入多道输出、多道输入和多道输出等系统。

根据算子的性质，系统分为线性系统和非线性系统两大类。线性系统算子一般用 L 表示，就是一个线性变换矩阵。

确定任意给定系统的特性很复杂，必须了解系统对任意输入的输出。若限定系统的一些特殊性质，如线性、时不变性、物理可实现性、稳定性、最小相位性等，就比较容易研究了。

（1）线性。对于任意输入信号 $x(t)$ 和任意常数 a，满足比例和叠加性质，即

$$\left.\begin{aligned} \sigma\{ax(t)\} &= a\sigma\{x(t)\} \\ \sigma\left\{\sum_{i=0}^{\infty} a_i x_i(t)\right\} &= \sum_{i=0}^{\infty} a_i \sigma\{x_i(t)\} \end{aligned}\right\} \tag{2-21}$$

其中，σ 为线性系统算子。

（2）时不变性。如果决定输入和输入的关系的系统特性不随时间变化，则称系统为时不变系统。如果系统由一个包含输入和输出的微分方程描述，时不变的含义是微分方程的系数与时间无关。用数学公式表示为

$$y(t-\tau) = \sigma\{x(t-\tau)\} \tag{2-22}$$

（3）稳定性。若一个系统对任意的有界输入，其输出都是有限的，称该系统稳定。

（4）物理可实现性。任何物理系统的输出在时间上不能产生与相应的输入之前。即若 $t < t_1$ 时 $x(t)=0$，则

$$y(t) = \sigma\{x(t)\} = 0, t < t_1 \tag{2-23}$$

三、输入与输出信号的关系

滤波器为线性时不变和物理可实现的稳定系统时，输入和输出之间的关系可用单位脉冲响应来描述。在系统的输入端，施加单位脉冲函数 $\delta(t)$。该单位脉冲函数可以看作具有无穷小宽度 $\mathrm{d}t$、脉冲下面积是 1 的矩形脉冲：

$$\int_{-\infty}^{\infty} \delta(t)\,\mathrm{d}t = 1 \tag{2-24}$$

在时间区间 dt 之外脉冲的幅值为零。用 $h(t)$ 表示系统对单位脉冲函数 $\delta(t)$ 的响应，即

$$\left.\begin{aligned} h(t) &= L\{\delta(t)\} \\ ah(t-\tau) &= L\{a\delta(t-\tau)\} \end{aligned}\right\} \tag{2-25}$$

其中，a 为比例常数；τ 为时间延迟。当任意函数 $x(t)$ 作为输入时，可认为输入是由很多窄的矩形脉冲组成，宽度为 dt。当 $x(t)$ 通过系统时，每个矩形脉冲都产生响应，输出为这些响应的叠加。如在时间 τ 时，脉冲宽度为 $d\tau$，高度为 $x(\tau)$，面积 $x(\tau)d\tau$，对应的 t 时刻输出的微分为

$$dy(t) = x(\tau)h(t-\tau)d\tau \tag{2-26}$$

则 t 时刻总的输出为

$$y(t) = \int_{-\infty}^{t} x(\tau)h(t-\tau)d\tau \tag{2-27}$$

右边的积分称为褶积或叠加积分，表示系统能用积分算子表示其特性。任何线性时不变系统的输入输出关系都由函数 $h(t)$ 给出，称为系统响应函数。线性时不变系统的输出为输入和系统响应函数的褶积。式(2-27)在信号与系统理论中极为重要。根据因果性，当 $\tau > t$ 时，$h(t-\tau)=0$，故式(2-27)可以改成

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau \tag{2-28}$$

利用变量替换 $\tau_1 = t-\tau$，故式(2-27)又可以写成

$$y(t) = \int_{-\infty}^{\infty} h(\tau_1)x(t-\tau_1)d\tau_1 \tag{2-29}$$

考虑因果性，$\tau_1 < 0$，$h(\tau_1)=0$，为方便，去掉下标，进一步写成

$$y(t) = \int_{0}^{\infty} h(\tau)x(t-\tau)d\tau \tag{2-30}$$

信号与系统在 Fourier 变换域的关系在后面给出。

信号通过滤波器，分几种情形：一种从已知输出和滤波器响应导出输入，称信号重构或反演；一种是根据输入和输出的关系来确定系统的响应函数，即所谓的系统辨识；还有仅仅知道输出，根据一些统计假设同时确定系统响应和输入，则为盲反褶积问题。

按照系统论的观点，包括前述的线性褶积模型问题在内，地震反褶积可以纳入一个统一的框架——盲系统辨识(Blind System Identification)。盲系统辨识是信号处理的一项基本技术，其目的是仅仅从系统(包括线性)的输出，估计未知输入和系统响应，相当于地震反射率函数和地震子波。图 2-1 为盲系统辨识示意图。习惯上也把这种与褶积模型类似的盲系统辨识称为盲反褶积。

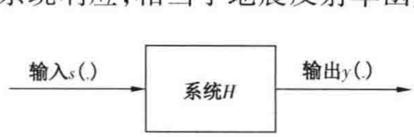


图 2-1 未知输入和系统响应的盲系统示意图

20 世纪 80 年代人们研究盲系统识别主要是单输入—单道—单输出线性系统情况，90 年代后，人们研究趋向于单道输入—多道—多输出线性系统。近年来，人们

对多输入—多输出线性系统模型进行了深入研究，正在和即将开展的是非线性系统情况的研究。地震勘探恰恰是多道数据采集系统，对于单炮道集，记录为单输入—多道—多输出的系统模型（多个系统函数）；对于共中心点道集或共像点道集记录为多输入—多输出系统模型（单个系统函数）。若为线性时不变系统，则为褶积模型；对于非线性系统，则为一般性系统辨识问题。信号处理界和地震界已经有大量研究文章并给出了很多有效的方法。图2－2给出反射地震勘探盲系统识别模型。

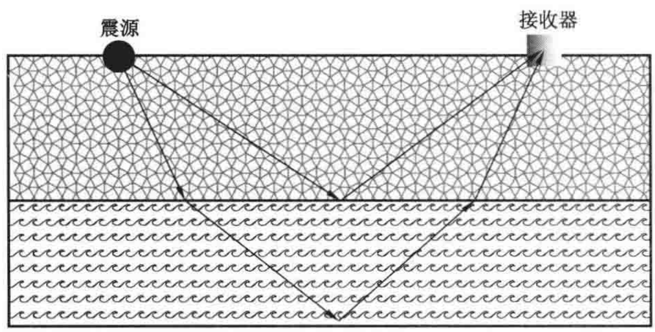


图2－2　反射地震勘探盲系统模型

地震勘探的信号与系统模型及估计与识别问题是地球物理信息处理当中的一个重要的内容。

四、离散时间信号处理

尽管地球物理测量的物理属性是模拟和连续的信号，但是在地球物理资料用计算机处理之前必须转换成数字和离散信号的形式。因此，离散信号的处理更为重要。关于采样和离散化将在后面进行详细讨论。离散时间信号处理用Z变换讨论。

第三节　离散随机信号基础

在实际中遇到的信号大多数信号是具有随机性质的信号，该随机信号可用随机过程来表示。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，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，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，但是一个随机信号服从确定的概率分布和联合概率分布，即对于指定时刻，信号取值的概率分布和概率密度函数是确定的，对一组指定的时间集合，信号各时刻取值变量之间的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），这些统计量反映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。

物理世界的大多随机信号是连续的信号，实际处理时，通过采样得到离散随机信号。

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。但在工程实际当中,可以获得的往往是随机过程的一次实现,如地震勘探中的数据采集,就是地震波信号在地层中传播的一次实现。统计信号处理的一个重要任务就是通过这样的一次实现,去估计随机信号的各种统计特征量及其这些统计特征量的各种变换形式,从而提取有用的信息。我们称离散随机信号的一次实现为一个时间序列,记作 $\{x(n)\}$,或简记为 $x(n)$,如每道地震信号就可看作一个时间序列。$\{x(n)\}$ 只表示随机信号的一次实现,且只是一个时间序列,随机信号则是它所有可能实现的集合。一个随机信号(过程)可以用 $\{x(n,\xi)\}$ 表示,ξ 表示随机过程的一次实验,$\{x(n,\xi)\}$ 表示所有实验的集合。为叙述简单,也用 $x(n)$ 表示一个随机信号。在随机信号中,各态遍历(历经)随机过程是指一次实现做出的时间统计特性和全体实现的时间统计特性相同,即全体实现在某时刻的状态等同于某次实现在所有时间里经历的状态。换句话,当时间序列长度趋于无穷时,时间平均与汇集平均相等。平稳随机过程,指随机过程的统计特性不随时间的推移而变化,即在时间 t_1 到时间 t_2 这段随机信号和时间 $t_1+\tau$ 到时间 $t_2+\tau$ 这段随机信号的统计特性基本相同,也就是说与起始时间无关。显然,一个具有各态遍历的随机过程一定是平稳的随机过程,而一个平稳的随机过程不一定是各态遍历的。在瞬变时间附近,随机过程表现为非平稳特性,需采用非平稳信号分析方法。

一、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的随机信号,时间序列是由一组按照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取时间的集合,给出它们的联合概率分布函数,一般用:

$$F(x_1,x_2,\cdots,x_M;n_1,n_2,\cdots,n_M)=P\{x(n_1)\leqslant x_1,x(n_2)\leqslant x_2,\cdots,x(n_M)\leqslant x_M\} \tag{2-31}$$

其中,$P\{\xi\}$ 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 $x_1,x_2,\cdots,x_M$ 可导,定义随机信号在时刻 $n_1,n_2,\cdots,n_M$ 取值的联合概率密度函数为

$$p(x_1,x_2,\cdots,x_M;n_1,n_2,\cdots,n_M)=\frac{\partial F(x_1,x_2,\cdots,x_M;n_1,n_2,\cdots,n_M)}{\partial x_1\partial x_2\cdots\partial x_M} \tag{2-32}$$

当

$$p(x_1,x_2,\cdots,x_M;n_1,n_2,\cdots,n_M)=p(x_1;n_1)p(x_2;n_2)\cdots p(x_M;n_M) \tag{2-33}$$

称随机信号在各个时刻取值统计上相互独立。若

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

则称随机信号在时间子集 $n_1,n_2,\cdots,n_l$ 的取值和时间子集 $n_l,n_{l+1},\cdots,n_M$ 的取值统计上独立,但在每个时间子集内不一定独立。

此为试读,需要完整PDF请访问: www.ertongbook.com