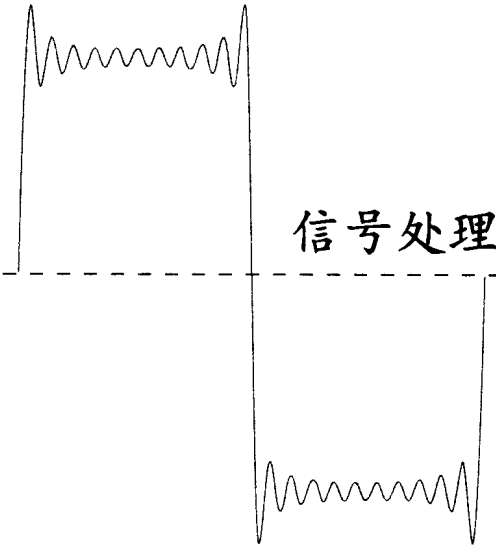


信号处理中的数学变换和估计方法

徐伯勋 白旭滨 傅孝毅 编著

清华大学出版社



信号处理中的数学变换和估计方法

徐伯勋 白旭滨 傅孝毅 编著

清华大学出版社
北京

内 容 简 介

本书介绍了信号处理中常用到的数学变换及估计方法,在写法上避开了繁杂的数学论证,努力做到深入浅出、通俗易懂.书中融入了作者多年来从事信号处理方面科研工作的实际经验.关于数学变换,既涉及傅里叶变换、Z变换、拉普拉斯变换、小波变换、希尔伯特变换、沃尔什变换、数论变换等正交变换,也涉及一类特殊的非线性变换——同态变换.关于估计方法,涉及了最小线性方差估计、最大似然估计、最小平方估计,以及现代谱估计方法.本书可作为应用数学、计算数学、地球物理数据处理和通信技术等专业的大学生和研究生参考书,也可供有关专业工程技术人员参考.

版权所有,翻印必究.举报电话:010-62782989 13901104297 13801310933

图书在版编目(CIP)数据

信号处理中的数学变换和估计方法/徐伯勋,白旭滨,傅孝毅编著. —北京:清华大学出版社,2004.7
ISBN 7-302-08250-2

I. 信… II. ①徐… ②白… ③傅… III. ①信号处理—变换 ②信号处理—估计 IV. TN911.7

中国版本图书馆 CIP 数据核字(2004)第 018445 号

出 版 者:清华大学出版社

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

社 总 机:010-62770175

地 址:北京清华大学学研大厦

邮 编:100084

客户服务:010-62776969

组稿编辑:刘 颖

文稿编辑:王海燕

印 刷 者:北京四季青印刷厂

装 订 者:三河市化甲屯小学装订二厂

发 行 者:新华书店总店北京发行所

开 本:185×230 印张:20 字数:409 千字

版 次:2004 年 7 月第 1 版 2004 年 7 月第 1 次印刷

书 号:ISBN 7-302-08250-2/O·350

印 数:1~3000

定 价:35.00 元

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换.联系电话:(010)62770175-3103 或(010)62795704

前言

本书部分内容是根据作者多年的科研工作和对信号处理的实际经验写成的. 在写法上对繁杂的数学论证和较熟悉的内容, 不做详细介绍. 尽量做到深入浅出、通俗易懂.

我们知道, 携带信息的物理过程称为信号, 它可表示成一个或几个独立自变量的函数. 一般常用一个以时间或空间为自变量的函数(图像信号为两个自变量的函数等)来表示. 它是传递信息的函数. 信号处理在数学上就是一种数学变换和估计方法. 本书试图介绍给读者尽可能多的数学变换和主要的滤波方法.

由于信号最基本的组合规则为加法、乘法和卷积这三种形式, 更复杂的信号组合规则可能就是这三种运算的复合形式. 书中讲到的数学变换主要是针对线性正交变换和一类特殊的非线性变换即同态变换(同态滤波或广义线性变换). 同态变换是专门对乘法和卷积组合而成的信号来考虑的. 在线性正交变换中最基本的变换为傅里叶变换及其有关的拉普拉斯变换、 Z 变换、窗口傅里叶变换、小波变换和希尔伯特变换等, 另外还介绍了一些其他的正交变换. 从某种意义上讲, 它们要比傅里叶变换更简单、方便, 也是在信号处理中较常用的方法, 如沃希变换、数论变换、哈尔变换和 $K-L$ 变换等. 在估计方法中, 我们主要介绍了最小线性方差估计、最大似然估计和最小平方估计, 以及维纳滤波、卡尔曼滤波等, 它们也属于最小线性方差估计之列, 且在信号处理中占有重要位置. 在谱估计中, 我们主要介绍了现代谱估计的方法、原理, 如最大熵谱估计, 最大似然谱估计和多谱估计, 特别对多谱估计在频域、时域上的计算方法做了较多的介绍.

全书共 8 章. 第 1 章为预备知识, 主要为读者介绍一些在信号处理中常用的内容; 第 2 章为信号处理中最基本的数学变换, 主要介绍傅里叶变换及与其有关的正交变换, 它们也是传统谱估计与滤波的基础; 第 3 章(沃尔什变换)主要介绍了三种不同定义的沃尔什变换, 以及它们之间的关系和在通信技术中的应用; 第 4 章(数论变换)介绍了数论变换的一些突出优点以及存在的问题; 第 5 章(其他正交变换)所介绍的正交变换都是在信号处理中有用的; 第 6 章(维纳滤波、伯格滤波和卡尔曼滤波)所介绍的滤波都是从搀杂噪声的信号中提取有用信号的常用而有效的滤波方法, 在科技和工程应用领域中都已得到了广泛应用; 第 7 章(同态滤波)主要解决以乘法和卷积组合而成的信号的分离(滤波)问题; 第 8 章(谱估计)重点介绍了现代谱估计中的三种方法.

清华大学工程力学系张如一教授审阅了全稿,并提出许多宝贵的意见和建议,在此谨致衷心的感谢。

本书可作为高等院校应用数学、计算数学、地球物理数据处理和通信技术等有关专业本科生及研究生的参考书,也可供从事数字信号处理的工程技术人员学习参考。

限于作者的水平和经验,书中肯定存在不少缺点和错误,殷切地希望读者批评指正。

作者

目 录

第 1 章 预备知识	1
1.1 n 维向量空间的概念	1
1.1.1 向量的范数(模)	2
1.1.2 l_2 空间与 L_2 空间的定义及其关系	4
1.1.3 两向量的内积	4
1.2 线性变换与矩阵	5
1.2.1 几种特殊矩阵	7
1.2.2 矩阵的奇异值分解	9
1.3 正交变换与正交函数系	10
1.3.1 正交函数系的封闭性与完备性	11
1.3.2 正交函数系与线性独立函数系的关系	12
1.4 信号的线性系统	14
1.4.1 信号的表达形式	14
1.4.2 离散信号	14
1.4.3 离散线性系统和滤波器分类	16
1.5 信号的一些基本知识	18
1.5.1 信号的最小、最大和混合延迟(相位)的定义	18
1.5.2 有限长度物理可实现信号的反信号	21
1.6 把给定信号转换成最小相位信号和零相位信号的方法	23
1.6.1 把已知信号转换成最小相位信号	23
1.6.2 把已知信号转换成零相位信号的几种方法	25
1.7 序列的卷积和相关	26
1.7.1 卷积	26
1.7.2 相关	30
1.7.3 卷积与相关之间的关系	33
1.8 Wold 定理——有限数据分解的理论基础	36
1.9 随机数字信号(序列)	38

2009/8/07

1.9.1	引言	38
1.9.2	随机序列的主要数字特征及其间的关系	38
1.10	随机信号(序列)的参数估计	43
1.10.1	引言	43
1.10.2	参数的估计	44
1.10.3	估计量的评价标准	48
1.10.4	参数估计和极小化	49
1.11	熵的概念及其性质	50
1.11.1	熵的概念	50
1.11.2	信息熵	51
1.11.3	熵的基本性质	52
第 2 章	信号处理中最基本的数学变换	54
2.1	引言	54
2.2	周期函数	54
2.2.1	周期函数的概念和性质	54
2.2.2	周期函数的傅里叶级数展开式	55
2.2.3	周期函数的频谱分析	57
2.3	非周期函数	58
2.3.1	非周期函数的傅里叶变换	58
2.3.2	傅里叶变换的性质	59
2.4	连续函数的卷积与相关	62
2.4.1	卷积定理	62
2.4.2	相关函数和能量谱密度的关系	63
2.5	拉普拉斯变换	63
2.5.1	引言	63
2.5.2	拉普拉斯变换的定义	64
2.5.3	拉普拉斯变换的性质	65
2.6	二维傅里叶级数与傅里叶变换	67
2.6.1	二维傅里叶级数	67
2.6.2	二维傅里叶变换	68
2.7	离散傅里叶变换	69
2.7.1	采样定理	69
2.7.2	离散傅里叶变换的定义	70

2.7.3	离散傅里叶变换的性质	71
2.7.4	二维离散傅里叶变换	72
2.8	Z 变换	73
2.8.1	Z 变换的定义	73
2.8.2	Z 变换的收敛域和惟一性问题	73
2.8.3	Z 变换与傅里叶变换的关系	74
2.8.4	Z 变换的性质	75
2.8.5	Z 逆变换	76
2.8.6	用 Z 变换解差分方程	78
2.9	窗口傅里叶变换、小波变换及其他变换	79
2.9.1	引言	79
2.9.2	窗口傅里叶变换	79
2.9.3	小波变换	80
2.9.4	声音变换	83
2.9.5	信号的小波分解和重建的例子	84
2.10	希尔伯特变换	85
第 3 章	沃尔什变换	90
3.1	引言	90
3.2	预备知识	90
3.3	第一类 按列率或沃尔什编号的沃尔什函数	93
3.3.1	定义	94
3.3.2	沃尔什函数的主要性质	97
3.3.3	沃尔什级数	99
3.3.4	采样定理及离散沃尔什变换	99
3.3.5	离散沃尔什变换的性质	103
3.3.6	沃尔什变换的快速算法	104
3.3.7	谱分析	106
3.4	第二类 按自然编号或阿达马编号的沃尔什函数	107
3.4.1	定义	107
3.4.2	第二类离散沃尔什变换的表示形式	109
3.4.3	第二类有限沃尔什变换的快速算法	110
3.5	第三类 按并元或佩利编号的沃尔什函数	111
3.5.1	定义	111

3.5.2	第三类有限沃尔什变换的表示形式	111
3.5.3	第三类有限沃尔什变换的快速算法	113
3.6	三类有限沃尔什变换之间的相互转换关系	114
3.7	二维有限沃尔什变换	116
3.7.1	二维沃尔什级数展开	117
3.7.2	二维有限沃尔什变换	117
3.8	沃尔什变换在通信技术中的应用	118
第4章	数论变换	121
4.1	引言	121
4.2	预备知识	121
4.3	具有循环卷积特性的变换结构	126
4.4	数论变换及存在定理	133
4.4.1	几种典型的数论变换	134
4.4.2	数论变换的性质	136
4.4.3	快速数论变换	144
4.4.4	数论变换中参数 M 、 N 和 α 的选择	147
4.4.5	用数论变换计算循环卷积	150
4.4.6	费马数变换	151
4.4.7	数论变换的应用	156
第5章	其他正交变换	159
5.1	哈尔变换	159
5.1.1	连续哈尔函数的定义	159
5.1.2	哈尔函数系的正交完备性	161
5.1.3	哈尔函数与沃尔什函数的关系	161
5.1.4	离散哈尔变换(DHT)	161
5.2	斜变换(ST)	163
5.2.1	引言	163
5.2.2	斜矩阵	163
5.2.3	斜变换的定义和快速算法	166
5.3	离散余弦变换(DCT)	167
5.3.1	离散余弦变换的定义	168
5.3.2	计算方法	169

5.4 正交变换在随机数字信号中的应用	170
5.4.1 引言	170
5.4.2 主成分分析法	170
5.4.3 主成分的主要性质	173
5.5 K-L 变换	173
5.5.1 K-L 变换的概念	173
5.5.2 K-L 变换的两个重要性质	174
5.5.3 举例	175
5.5.4 应用——K-L 滤波	177
5.6 奇异值分解	178
5.6.1 矩阵的奇异值与奇异值分解	178
5.6.2 奇异值分解法的优缺点	180
5.6.3 奇异值分解和 K-L 变换之间的关系	180
5.6.4 奇异值分解在垂直地震剖面中的应用	181
5.7 灰色系统中的一种数学变换	184
5.7.1 概述	184
5.7.2 具体实现步骤	184
第 6 章 维纳滤波、伯格滤波与卡尔曼滤波	187
6.1 维纳滤波	187
6.1.1 单道维纳滤波的数学模型	187
6.1.2 维纳滤波因子的求法	187
6.1.3 维纳滤波因子的实际求法	188
6.1.4 最小平方反卷积	189
6.1.5 预测反卷积	193
6.1.6 波形反卷积	198
6.1.7 维纳滤波(最小平方滤波)的有关性质	199
6.1.8 多道维纳滤波	206
6.2 伯格滤波(反卷积)	210
6.2.1 引言	210
6.2.2 $AR(M)$ 模型等价于一步预测反卷积	211
6.2.3 由 $AR(M)$ 出发如何外推自相关值	211
6.2.4 $AR(M)$ 模型与最大熵外推自相关序列是等价的	212

6.2.5	怎样从已知的 $r_{xx}(0), r_{xx}(1), \dots, r_{xx}(M)$ 去计算 $a_{M,m} (m=0, 1, \dots, M)$	214
6.2.6	伯格算法(最大熵法)	215
6.2.7	向前预测误差和向后预测误差的滤波作用	217
6.2.8	应用——最大熵子波反卷积	218
6.3	卡尔曼滤波	221
6.3.1	引言	221
6.3.2	一维滤波	223
6.3.3	多维滤波	224
6.3.4	线性离散系统的卡尔曼滤波公式的推导	226
6.3.5	总结	229
第7章	同态滤波	231
7.1	信号组合规则的三种最基本形式	231
7.2	线性滤波器与非线性滤波器	232
7.2.1	线性滤波器(线性系统)	232
7.2.2	一类特殊的非线性滤波器(同态滤波)	234
7.3	卷积型同态滤波	236
7.3.1	卷积型同态滤波的标准形式	236
7.3.2	输入特征系统 D	236
7.3.3	复赛谱的性质和求法	239
7.3.4	线性系统 L	251
7.3.5	输出特征系统 D^{-1}	253
7.3.6	同态滤波的应用——反卷积和反鸣震	255
第8章	谱估计	260
8.1	引言	260
8.2	确定性信号的谱估计	261
8.3	平稳随机信号的谱估计	263
8.4	传统谱估计方法	264
8.4.1	直接法	264
8.4.2	间接法	265
8.5	现代谱估计方法	265
8.5.1	最大熵谱估计法	265

8.5.2 最大熵谱估计中振幅谱和相位谱的求法.....	273
8.5.3 最大似然谱.....	277
8.5.4 多谱估计.....	279
8.5.5 多谱估计的频域方法.....	280
8.5.6 多谱估计的时域方法——参数方法.....	284
8.6 最大熵谱估计在油气检测中的应用	291
参考文献	299
索引	302

第1章 预备知识

1.1 n 维向量空间的概念

定义 1.1 由 n 个有顺序的实数组 $(x_1, x_2, \dots, x_n)$ 的全体构成的集合称做 n 维空间. 实数组 $(x_1, x_2, \dots, x_n)$ 称为 n 维空间的一点; 而 $x_1, x_2, \dots, x_n$, 则称为这个数组的分量或坐标.

由于点 $(x_1, x_2, \dots, x_n)$ 又可代表从原点 $(0, 0, \dots, 0)$ 出发的以点 $(x_1, x_2, \dots, x_n)$ 为终点的向量, 在代数上常把数组 $(x_1, x_2, \dots, x_n)$ 称为 n 维空间的一个 n 维向量, n 维空间又称 n 维向量空间.

显然, 当 $n=1$ 时, 一维向量空间就代表一条直线. 只要在直线上取定一个原点, 取定一个方向和单位长度, 建立坐标系后, 则直线上的任意一点 x 便与一个实数一一对应了. 这样, 把全体实数的集合认为是一维向量空间.

当 $n=2$ 时, 在平面(二维向量空间)上建立一个坐标系后, 平面上任一点就与一对有序数组 (x_1, x_2) 一一对应, 不同的数组代表不同的点, 顺序变了, 点也不同了. 数组 (x_1, x_2) 为一个二维向量.

当 $n=3$ 时, 在三维空间建立坐标系后, 任一点就与它的坐标 (x_1, x_2, x_3) 一一对应, 这时坐标 (x_1, x_2, x_3) 是三个有顺序的实数组, 也称三维向量. 当 $n>3$ 时, 就没有几何直观了.

定理 1.1 设 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 是 n 个线性无关的 n 维向量, 而 x 是任意的 n 维向量, 则 x 一定能表示成 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 的线性组合.

证明 因为 $n+1$ 个 n 维向量必定线性相关, 即

$$lx + a_1\varepsilon_1 + a_2\varepsilon_2 + \dots + a_n\varepsilon_n = 0.$$

因 $l \neq 0$, 所以

$$\begin{aligned} x &= \left(-\frac{a_1}{l}\right)\varepsilon_1 + \left(-\frac{a_2}{l}\right)\varepsilon_2 + \dots + \left(-\frac{a_n}{l}\right)\varepsilon_n \\ &= x_1\varepsilon_1 + x_2\varepsilon_2 + \dots + x_n\varepsilon_n, \end{aligned}$$

其中 $x_i = -\frac{a_i}{l} (i=1, \dots, n)$.

上述结果表示在 n 维向量空间中, 任一 n 维向量 x 都可以表示成 n 个线性无关的向量 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 的线性组合. 且这种表示由其系数 $x_1, x_2, \dots, x_n$ 惟一决定.

上述概念其实在解析几何中早已遇到过. 如在二维空间(平面)上任一向量可以表示成2个不共线的向量的线性组合; 在三维空间上任一向量都可以表示成3个不共面的向量的线性组合. 特别在二维空间中取两个互相垂直的单位向量 $i=(1,0)$, $j=(0,1)$, 在三维空间中取3个互相垂直的单位向量 $i=(1,0,0)$, $j=(0,1,0)$, $k=(0,0,1)$; 就分别组成了平面上和空间中直角坐标系的坐标向量. 这样, 定理1.1也可以理解为任一个向量可沿坐标向量进行分解, n 维空间中任意 n 个线性无关的向量都有着与坐标向量类似的性质.

定义 1.2 n 维空间中的一组 n 个线性无关的向量叫做 n 维空间的一组基或基向量. 基向量就是坐标向量的推广.

1.1.1 向量的范数(模)

在欧几里得空间中, 设 a, b 为任意数, 我们常用绝对值 $|a|$, $|b|$ 和 $|a-b|$ 来衡量 a, b 的大小及 a 与 b 相差的大小. 在微积分学中的极限论中, 我们也常用长度(绝对值)来刻画收敛的性质, 如 $\lim_{n \rightarrow \infty} x_n = x_0$, 常用当 $n \rightarrow \infty$ 时 $|x_n - x_0| \rightarrow 0$ 来刻画. 在数据处理中, 涉及收敛性、稳定性和误差分析等方面, 都需要衡量两个向量 x, y 本身以及它们差的大小, 我们也应构造适当的概念来反映它们, 这就是要介绍的向量范数的概念. 它是比长度意义更为广泛的概念, 其定义如下.

定义 1.3 如果 V 是数域 K 上的线性空间, 且对于 V 的任一向量 x , 对应一个实值函数 $\|x\|$, 它满足以下3个条件:

1. 非负性 当 $x \neq 0$ 时, $\|x\| > 0$; 当 $x = 0$ 时, $\|x\| = 0$;
2. 齐次性 $\|ax\| = |a| \|x\|$, $a \in K, x \in V$;
3. 三角不等式 $\|x+y\| \leq \|x\| + \|y\|$, $x, y \in V$.

则称 $\|x\|$ 为 V 上向量 x 的范数, 简称向量范数.

最常见的向量范数有3种:

$$\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2};$$

$$\|x\|_\infty = \max_i |x_i|;$$

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|.$$

可以证明, 它们是符合上述3个条件的.

称 $\|x\|_2$ 为 2-范数或欧几里得范数, 它对欧几里得空间 $\mathbb{R}^n$ 亦成立. 只要把复数域 $\mathbb{C}$ 改为实数域 $\mathbb{R}$ 就行了, 此时 $\|x\|_2$ 就是通常所说的长度或距离.

可以证明范数 $\|x\|_2$ 还满足不等式

$$|\|x\|_2 - \|y\|_2| \leq \|x - y\|_2, \quad (1.1)$$

这里 x, y 的 $\mathbb{C}^n$ 的任意向量.

如用 $-y$ 来代替(1.1)式中的 y , 不等式(1.1)成为

$$|\|x\|_2 - \|y\|_2| \leq \|x+y\|_2. \quad (1.2)$$

不等式(1.1)和不等式(1.2)在 $\mathbb{R}^2$ (平面几何)中有明确的几何意义,即表示任一三角形两边长度之差不大于第三边的长度(见图 1.1).

类似地,我们称 $\|x\|_\infty$ 为 ∞ -范数,称 $\|x\|_1 = \sum_{i=1}^n |x_i|$ 为1-范数.

我们称函数 $(\sum_{i=1}^n |x_i|^p)^{1/p}$ 为向量 x 的 p -范数或 l_p 范数,记为 $\|x\|_p$. 于是有

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}. \quad (1.3)$$

在(1.3)式中,当 $p=1$ 时,便为 $\|x\|_1$;当 $p=2$ 时为 $\|x\|_2$;并且有

$$\|x\|_\infty = \lim_{p \rightarrow +\infty} \|x\|_p.$$

这是因为,如 $|x_1|, |x_2|, \dots, |x_n|$ 中最大的一个是 $|x_{iL}| \neq 0$,则有

$$\|x\|_\infty = \max_i |x_i| = |x_{iL}|.$$

于是

$$\|x\|_p = \left(\sum_{i=1}^n |x_{iL}|^p \frac{|x_i|^p}{|x_{iL}|^p} \right)^{1/p} = |x_{iL}| \left(\sum_{i=1}^n \frac{|x_i|^p}{|x_{iL}|^p} \right)^{1/p},$$

又因为 $|x_{iL}|^p \leq \sum_{i=1}^n |x_i|^p \leq n |x_{iL}|^p$,所以有

$$1 \leq \left(\sum_{i=1}^n \frac{|x_i|^p}{|x_{iL}|^p} \right)^{1/p} \leq n^{1/p},$$

由洛必达法则可知 $\lim_{p \rightarrow +\infty} n^{1/p} = 1$. 故有

$$\lim_{p \rightarrow +\infty} \|x\|_p = |x_{iL}| = \|x\|_\infty.$$

我们称(1.3)式为 l_p 范数(模). 由于 $p(1 \leq p < +\infty)$ 是任意实数,所以 l_p 包括了无限多种范数,我们称常见的 $\|x\|_1, \|x\|_2, \|x\|_\infty$ 为 l_1, l_2, l_∞ 范数(也称模).

显然,对于一个向量,定义不同的范数,其大小也可不同,见下例.

例 1.1 计算 $\mathbb{C}^4$ 的向量 $x = (3i, 0, -4i, -12)$ 的 l_1, l_2, l_∞ 范数.

解

$$\|x\|_1 = \sum_{i=1}^4 |x_i| = |3i| + |0| + |-4i| + |-12| = 19,$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^4 |x_i|^2} = \sqrt{x \bar{x}} = \sqrt{(3i)(-3i) + (-4i)(4i) + (-12)^2} = 13,$$

$$\|x\|_\infty = \max_4 (|3i|, |0|, |-4i|, |-12|) = 12.$$

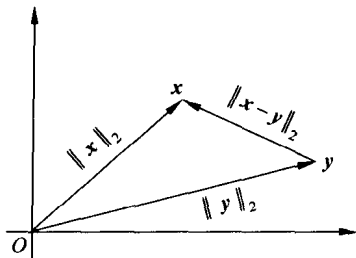


图 1.1 三角形两边之差
不大于第三边

1.1.2 l_2 空间与 L_2 空间的定义及其关系

定义 1.4 有序的实数序列 $x = (x_1, x_2, \dots, x_n, \dots)$ 如果满足 $\|x\| = \sqrt{\sum_{n=1}^{\infty} x_n^2} < \infty$

时, 则称 x 是 l_2 空间中的一个向量, 称 $\|x\|$ 为 x 的范数.

定义 1.5 满足条件

$$\int_a^b |f(x)|^p dx < \infty$$

的所有函数 $f(x)$ 的全体记作 $L_p(a, b)$, 其中 $f(x)$ 为复值函数.

特别地, 满足能量有限:

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt < \infty$$

的所有复值函数 $f(t)$ 的全体记作 $L_2(-\infty, \infty)$, 即能量有限的时间函数的全体构成 L_2 空间.

l_2 中每一个向量是 L_2 中某一个函数的傅里叶系数所组成的数列. 所以空间 l_2 和 L_2 有着密切的关系.

1.1.3 两向量的内积

设两向量 x, y 的夹角为 θ , 我们定义

$$|x| |y| \cos \theta$$

为 x 与 y 的内积, 记作 (x, y) , 即

$$(x, y) = |x| |y| \cos \theta. \quad (1.4)$$

由定义显然有

$$(x, y) = (y, x), \quad (x, x) = |x|^2.$$

若 x 与 y 互相正交(垂直), 则 $\cos \theta = 0$, 即 $(x, y) = 0$; 反之若 x 和 y 都不是零向量, 而 $(x, y) = 0$, 则 $\cos \theta = 0$, 即 $\theta = 90^\circ$. 所以两向量正交(或垂直)的充分必要条件是它们的内积等于零. 这样, 如果两个三维向量 $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$ 正交, 即

$$(x, y) = |x| |y| \cos \theta = 0, \quad (1.5)$$

按解析几何的写法, 显然有

$$(x, y) = (x_1 i + x_2 j + x_3 k)(y_1 i + y_2 j + y_3 k) = x_1 y_1 + x_2 y_2 + x_3 y_3 = 0. \quad (1.6)$$

这是因为 $(i, i) = (j, j) = (k, k) = 1$, $(i, j) = (j, k) = (k, i) = 0$. 当 $x = y$ 时, $(x, x) = x_1^2 + x_2^2 + x_3^2$, 这里 i, j, k 为单位坐标向量, 它们是互相正交且不相干的. 有了正交性质, 向量的系数 x_1, x_2, x_3 就很容易求得, 即由 $x = x_1 i + x_2 j + x_3 k$ 得 $x_1 = (x, i)$, $x_2 = (x, j)$, $x_3 = (x, k)$.

推广到 n 维向量, 显然(1.6)式为

$$\sum_{i=1}^n x_i y_i = 0. \quad (1.7)$$

如设 n 维基向量为 $e_1, e_2, \dots, e_n$, 则 $x = \sum_{k=1}^n x_k e_k$, 其中 $x_k = (x, e_k)$; $e_k = (0, \dots, 0, 1_k, 0, \dots, 0)$, 1_k 表示第 k 个元素为 1.

如果我们将两向量的正交性推广到函数上, 很自然地, 会想到两个定义在区间 $[a, b]$ 上的函数 $u_1(t), u_2(t)$ 的正交性定义为

$$\int_a^b u_1(t) u_2(t) dt = 0, \quad (1.8)$$

区间 $[a, b]$ 称为正交区间.

1.2 线性变换与矩阵

我们知道, 空间维数是构成该空间线性无关向量组所含向量的最大个数. 如三维空间中, 可以有两个线性无关的向量, 也可以有三个线性无关的向量, 但空间的任意四个向量必定线性相关的. 我们还知道 n 维向量可以相加, 也可以数乘, 其结果还是 n 维空间的向量, 这称为 n 维空间对加法和数乘是封闭的. 因此任意一个 n 维向量 $x = (x_1, x_2, \dots, x_n)$ 可分解为

$$\begin{aligned} x &= (x_1, x_2, \dots, x_n) = (x_1, 0, \dots, 0) + (0, x_2, 0, \dots, 0) + \dots + (0, \dots, 0, x_n) \\ &= x_1(1, 0, \dots, 0) + x_2(0, 1, 0, \dots, 0) + \dots + x_n(0, 0, \dots, 0, 1) \\ &= x_1 e_1 + x_2 e_2 + \dots + x_n e_n, \end{aligned}$$

其中 $e_1, e_2, \dots, e_n$ 为 n 维空间的坐标向量(或基向量); $x_1, x_2, \dots, x_n$ 为坐标.

下面我们在三维空间里引入线性变换与矩阵的概念, 设 e_1, e_2, e_3 是一组基向量, 设向量 v 以 e_1, e_2, e_3 为基向量, 坐标是 y_1, y_2, y_3 , 即

$$v = y_1 e_1 + y_2 e_2 + y_3 e_3. \quad (1.9)$$

若另取一组基向量为 e'_1, e'_2, e'_3 , 这里 e'_1, e'_2, e'_3 在 e_1, e_2, e_3 的坐标依次为 (a_{11}, a_{21}, a_{31}) , (a_{12}, a_{22}, a_{32}) , (a_{13}, a_{23}, a_{33}) 即

$$\begin{cases} e'_1 = a_{11} e_1 + a_{21} e_2 + a_{31} e_3, \\ e'_2 = a_{12} e_1 + a_{22} e_2 + a_{32} e_3, \\ e'_3 = a_{13} e_1 + a_{23} e_2 + a_{33} e_3. \end{cases} \quad (1.10)$$

现在我们要问向量 v 在基向量 e'_1, e'_2, e'_3 下的坐标与在基向量 e_1, e_2, e_3 下坐标有何关系?

设 v 在基向量 e'_1, e'_2, e'_3 下的坐标为 x_1, x_2, x_3 , 即

$$v = x_1 e'_1 + x_2 e'_2 + x_3 e'_3.$$

由(1.9)式有