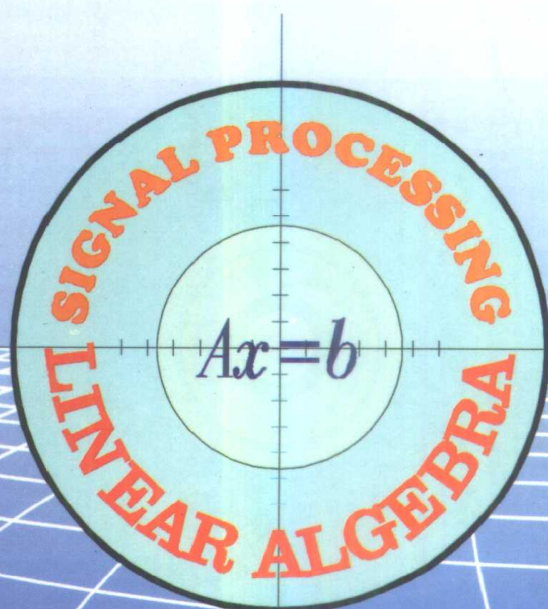


信号处理中的线性代数

张贤达 著



SVD TLS IV Toeplitz Hankel Hilbert ...

科学出版社



信号处理中的线性代数

张贤达 著

科学出版社

1997

内 容 简 介

本书从工程科学和技术角度出发,以信号处理应用为例,系统、深入地阐述线性代数的理论、方法与应用.全书以矩阵方程组 $Ax=b$ 的求解作为主线,共分 11 章,内容包括以下四部分:(1) 矩阵代数基础,特殊矩阵和矩阵分解;(2) 向量空间,特征子空间分析方法,子空间跟踪与更新;(3) 奇异值分解及其各种推广;(4) 总体最小二乘及其推广,各种推广的最小二乘和辅助变量方法.为方便读者学习,书中给出了大量高度概括重要理论和方法的具体算法,以及参考和引用过的近 300 篇重要文献,书后给出了较为完整的 700 余条索引.

本书可作为高等学校信号与信息处理、电子与电机工程、计算机、自动控制、通信、航空、航天、土木、应用数学、力学等专业的高年级大学生和研究生的教材或参考书,也可供从事有关研究与应用的广大科技人员学习与参考.

图书在版编目(CIP)数据

信号处理中的线性代数 / 张贤达著. —北京:科学出版社, 1997

ISBN 7-03-005548-9

I. 信… II. 张… III. 线性代数 - 应用 - 信号处理 IV. TN911.7

中国版本图书馆 CIP 数据核字 (96) 第 25679 号

科学出版社出版

北京东黄城根北街16号

邮政编码: 100717

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1997 年 7 月第 一 版	开本: 787 × 1092 1/16
1997 年 7 月第一次印刷	印张: 28 1/2
印数: 1 - 3 000	字数: 658 000

定价: 40.00 元

前 言

很久以来,线性代数不仅是各数学学科的基本工具,而且也是许多理工学科的主要数学工具.就其本身的研究来说,线性代数和矩阵理论也是极富创造性的领域.它们的创造性极大地推动和丰富了其它众多学科的发展.例如,在信号和信息处理、系统工程和控制理论中,许多新的理论、方法和技术的诞生与发展就是线性代数和矩阵理论的创造性应用与推广的结果.可以说,线性代数在物理、力学、信号与信息处理、系统、控制、土木、电子、通信、电机、航空和航天等众多学科中是最富创造性和灵活性,并起最重要作用的数学工具.

虽然国内出版了不少深受读者喜爱的线性代数和矩阵理论的书籍,但从数学角度著述的书往往内容的叙述和证明太抽象,非数学专业的读者难于读懂它们;而那些以工科读者为对象的书又显得内容过于单薄,远不能够满足较深层次读者的需要.另一方面,线性代数和矩阵理论近 10 年内许多重要的新理论、新方法和新应用也尚未在国内外系统成书.

作者在《现代信号处理》(清华大学出版社,1995 年)一书中曾有意地重点介绍了广义特征值分解、总体最小二乘、正交投影矩阵等有关理论与应用,并在近几年的教学中系统地讲授了这些内容,从而引起了广大研究生对这些线性代数方法的浓厚兴趣.但是,由于受题材和篇幅的限制,在该书中忍痛割爱了有关的大量重要内容,使许多对线性代数和矩阵理论需求比较多和比较深的研究生尤其是博士研究生感到很大的不满足.这种强烈的需求萌发了我斗胆另写一书专门介绍信号处理学科中的线性代数的念头.我的这一想法得到了一些信号处理专家和同仁特别是几位学术前辈的充分肯定与支持,他们还建议书名取作《信号处理中的线性代数》,并希望此书对信号处理学科以外的读者也具有比较大的参考和使用价值.

由于工程科学上的问题往往归结为矩阵方程组 $Ax = b$ 的求解,所以作者以此作为贯穿全书的主线,意图包括各类求解方法,特别是一些最新的方法.全书共分十一章,其主要内容可概括为

1. 矩阵代数基础、特殊矩阵和矩阵的分解 (第一至第四章);
2. 子空间理论: 向量空间、特征子空间分析方法、子空间跟踪与更新 (第五、十、十一章);
3. 奇异值分解及其各种推广 (第六章);
4. 在有扰动和噪声的情况下求解矩阵方程 $Ax = b$ 的各类方法: 总体最小二乘及其各种推广 (第七章), 极大似然法和各种推广的最小二乘方法 (第八章), 辅助变量方法 (第九章).

作者试图在以下几方面形成本书的特点:

1. 完全从工程科学角度著述, 重点阐明线性代数理论与方法的基本思想、灵活性和实用性. 作者相信, 这种将数学和工程应用有机结合的做法会有利于工程界的广大读者增加学习和应用线性代数的热情与积极性, 并将有关工程研究与应用深入到更高、更新

的层次,同时也将有助于数学界的读者了解很多生动的应用背景,引发他们对线性代数和矩阵理论应用研究的新兴趣.

2. 加大选材的广度和深度,充分体现内容的新颖性、先进性和实用性.为此,本书将近几年国际著名杂志上发表的 100 多篇文献的基本思想和主要方法加以分类、归纳和融合,首次系统成书.书中还介绍了作者自己如何应用线性代数灵活解决信号处理中的典型或重要问题的大量研究成果,它们主要发表在 IEEE 汇刊上.

3. 应用背景明确.书中介绍的线性代数和矩阵理论的重要理论和方法都有信号处理的例子作为引出和应用说明,而且大多数是创造性应用和发展线性代数的生动例子(如第五章的 5.6 和 5.7 节、第六至第十一章).有必要指出的是,这些丰富例子都用比较浅显的方式叙述,其它领域和学科的读者也容易看得懂;而在所介绍的重要理论和方法中,奇异值分解、总体最小二乘、辅助变量方法和特征子空间已成为信号处理和其它一些工程学科中的常用“语言”,因此具有广泛的代表性和适用性.

4. 书中给出了大量高度概括重要理论和方法的具体算法,列出了全书参考和引用过的近 300 篇文献,还汇编了 700 余条索引.这些算法、参考文献和索引从不同侧面提供了有关线性代数新成果和新发展的丰富信息,可以方便读者在教学、科研和技术工作中检索、参考与引用.

著述一部线性代数的书对于一个非数学专业出身的作者来说,无疑是一项艰巨的挑战.作者虽为此竭尽全力,但囿于水平和能力,书中未能如愿乃至不妥和错误之处可能不乏其例.在此,诚恳希望得到诸位专家、同仁和广大读者的批评指正.

作者在完成本书的过程中,承蒙中国科学院院士李衍达教授和保铮教授、北京大学数学所程乾生教授、清华大学应用数学系蔡大用教授、自动化系金以慧教授、北京航空航天大学电子工程系毛士艺教授和北京邮电大学诸维明教授给予热情关心、支持和帮助,他们为本书提出了诸多宝贵意见与建议,谨向他们表示诚挚谢意.

作者近几年有关现代信号处理的研究项目得到了下列基金的资助:国家自然科学基金、国家教委留学回国人员专项基金、高等学校博士点专项研究基金、航空科学基金、航天基础性研究基金和清华大学基础研究基金.综合研究项目“关于 ARMA 模型辨识与谐波恢复的研究”获 1996 年国家教委科技进步(甲类)一等奖,其中的大部分学术成果与线性代数的灵活和创新应用密切相关,它们在本书中得到了充分的反映.在此,对以上各基金的资助一并致谢.

张 贤 达

1996 年 9 月于清华大学

目 录

第一章 广义逆矩阵与 Kronecker 积	1
1.1 基本概念与符号	1
1.1.1 矩阵符号与基本矩阵运算	1
1.1.2 独立性、子空间、基与维数	5
1.1.3 值域、零空间与秩	5
1.1.4 向量内积与外积	6
1.2 范数	7
1.2.1 向量范数	7
1.2.2 向量范数作 Lyapunov 函数	8
1.2.3 矩阵范数	10
1.2.4 Hankel 算子的范数	11
1.3 逆矩阵	13
1.3.1 逆矩阵	13
1.3.2 矩阵求逆引理	14
1.4 特征值问题与广义特征值问题	16
1.4.1 特征值问题	16
1.4.2 广义特征值问题	19
1.5 广义逆矩阵	20
1.5.1 Moore-Penrose 逆矩阵	20
1.5.2 最小二乘解	21
1.5.3 最小范数解	22
1.5.4 广义逆矩阵的阶数递推计算	23
1.5.5 超定二维超越方程的求解	25
1.6 Kronecker 积	27
1.6.1 Kronecker 积及其性质	27
1.6.2 Kronecker 积的应用	31
参考文献	34
第二章 特殊矩阵	36
2.1 对称矩阵与循环矩阵	36
2.2 交换矩阵与置换矩阵	40
2.3 正交矩阵与酉矩阵	42
2.4 Hermitian 矩阵	43
2.5 带型矩阵	47
2.6 Vandermonde 矩阵	48

2.7 Hankel 矩阵	52
参考文献	55
第三章 矩阵的变换与分解	56
3.1 正交投影	56
3.2 Householder 变换	57
3.2.1 保范数性与协方差不变	58
3.2.2 Householder 变换算法	60
3.3 Givens 旋转	62
3.3.1 Givens 旋转	63
3.3.2 快速 Givens 旋转	64
3.3.3 Kogbetliantz 算法	66
3.4 相似变换与矩阵的标准型	67
3.4.1 相似变换	67
3.4.2 矩阵的标准型	68
3.5 矩阵分解的分类	69
3.6 对角化分解	70
3.7 Cholesky 分解与 LU 分解	72
3.7.1 Cholesky 分解	72
3.7.2 LU 分解	73
3.8 QR 分解及其应用	75
3.8.1 QR 分解的性质	75
3.8.2 采用修正 Gram-Schmidt 法的 QR 分解	76
3.8.3 采用 Householder 变换的 QR 分解	78
3.8.4 采用 Givens 旋转的 QR 分解	78
3.8.5 基于 QR 分解的参数估计问题	79
3.8.6 基于 Householder 变换的快速时变参数估计	81
3.8.7 基于 Givens 旋转的时变参数估计	83
3.9 三角-对角化分解	85
3.9.1 LDM^T 和 LDL^T 分解	85
3.9.2 相似变换	87
3.9.3 Schur 分解	88
3.10 三对角化分解	90
3.11 矩阵束的分解	92
参考文献	93
第四章 Toeplitz 矩阵	95
4.1 半正定性	95
4.2 特征值与特征向量	96
4.3 Toeplitz 线性方程组的 Levinson 递推求解	99
4.3.1 经典 Levinson 递推	100

4.3.2 分基 Levinson 算法	101
4.3.3 分基 Schur 算法	106
4.3.4 Hermitian Levinson 递推	107
4.3.5 多信道 Toeplitz 线性方程组的 Levinson 递推求解	111
4.4 求解 Toeplitz 线性方程组的快速算法	112
4.4.1 循环镶嵌	112
4.4.2 Toeplitz 矩阵的部分求逆	113
4.4.3 Toeplitz 线性方程组求解	119
4.5 Toeplitz 矩阵的快速余弦变换	121
4.5.1 Toeplitz 矩阵的快速余弦变换	121
4.5.2 应用	124
参考文献	125
第五章 向量空间	126
5.1 内积空间及其性质	126
5.2 Hilbert 空间	129
5.2.1 n 维 Hilbert 向量空间	129
5.2.2 无限维 Hilbert 向量空间	131
5.2.3 L_2 空间	132
5.3 投影定理与均方估计	132
5.3.1 投影定理	133
5.3.2 均方估计	134
5.4 新息过程与 Kalman 滤波	138
5.4.1 新息定理	138
5.4.2 Kalman 滤波	139
5.5 正交集与正交基	141
5.6 正交投影矩阵及其应用	144
5.6.1 投影矩阵和正交投影矩阵	144
5.6.2 更新公式	146
5.6.3 利用正交投影矩阵设计 LS 格型滤波器	149
5.6.4 投影矩阵的导数	152
5.7 横向滤波器算子及其应用	153
5.7.1 横向滤波器算子及其递推	153
5.7.2 快速横向滤波器更新	155
参考文献	159
第六章 奇异值分解	160
6.1 数值稳定性与条件数	160
6.2 奇异值分解	163
6.2.1 奇异值分解及其几何意义	163
6.2.2 奇异值的性质	165

6.2.3 秩亏缺最小二乘解	168
6.2.4 奇异值分解的数值计算	172
6.3 乘积奇异值分解	173
6.3.1 三角矩阵的奇异值分解	175
6.3.2 矩阵乘积的奇异值分解	176
6.3.3 乘积奇异值分解算法的实现	177
6.4 广义奇异值分解	179
6.4.1 对称正定问题	179
6.4.2 广义奇异值分解	181
6.4.3 广义奇异值分解的实际算法	186
6.4.4 二次型不等式约束最小二乘	189
6.5 约束奇异值分解	191
6.5.1 约束奇异值	192
6.5.2 约束奇异值分解	193
6.6 结构奇异值	198
6.6.1 结构奇异值的定义与性质	198
6.6.2 结构奇异值的计算	199
6.7 奇异值分解的应用	202
6.7.1 静态系统的奇异值分解	202
6.7.2 系统辨识	205
6.7.3 阶数确定	206
6.7.4 系统的可控性	209
6.8 广义奇异值分解的应用	210
参考文献	212
第七章 总体最小二乘方法	215
7.1 最小二乘方法	215
7.1.1 矩阵方程解的可辨识性	215
7.1.2 Gauss-Markov 定理	217
7.2 总体最小二乘: 理论与方法	219
7.2.1 总体最小二乘解	219
7.2.2 总体最小二乘解的性能	223
7.3 总体最小二乘: 应用	228
7.3.1 ARMA 建模的总体最小二乘法	228
7.3.2 频率估计的总体最小二乘法	229
7.3.3 FIR 自适应滤波的总体最小二乘算法	234
7.4 约束总体最小二乘	236
7.4.1 约束总体最小二乘方法	237
7.4.2 约束总体最小二乘与极大似然的关系	241
7.4.3 约束总体最小二乘解的扰动分析	243

7.4.4 应用	245
7.5 结构总体最小二乘	247
7.5.1 结构总体最小二乘解	247
7.5.2 结构总体最小二乘解的性质	250
7.5.3 迭代算法	251
7.5.4 秩亏缺 Hankel 矩阵逼近	252
7.5.5 有噪声的实现问题	253
7.6 全局总体最小二乘	256
7.6.1 静态总体最小二乘	257
7.6.2 全局总体最小二乘	258
7.6.3 状态表示	259
7.6.4 系统的最优逼近	262
7.6.5 最优性条件	263
参考文献	265
第八章 极大似然法与推广的最小二乘方法	268
8.1 极大似然法	268
8.1.1 极大似然准则	268
8.1.2 基于特征结构的极大似然估计子	270
8.1.3 迭代二次型极大似然 (IQML) 算法	273
8.2 广义最小二乘方法	275
8.3 渐近最小方差估计	279
8.4 加权最小二乘方法	283
8.4.1 最优加权最小二乘估计	284
8.4.2 渐近最优加权最小二乘估计	285
参考文献	288
第九章 辅助变量方法	289
9.1 基本的辅助变量方法	289
9.1.1 基本辅助变量方法	290
9.1.2 辅助变量的选择方法	291
9.2 最优辅助变量方法	296
9.2.1 扩展的辅助变量方法	296
9.2.2 最优辅助变量估计	297
9.2.3 一致性与精度分析	298
9.2.4 $Q(q^{-1})$ 的最优选择	300
9.3 超定的递推辅助变量方法	301
9.3.1 超定的递推辅助变量方法	301
9.3.2 双曲变换	304
9.3.3 平方根超定递推辅助变量算法	307
9.4 阶数递推的辅助变量方法	311

9.5 辅助变量方法在模型阶数确定中的应用	314
9.5.1 最小描述长度 (MDL) 准则	314
9.5.2 综合 MDL 准则与辅助变量方法的模型定阶	316
9.5.3 与奇异值分解定阶方法的关系	318
参考文献	319
第十章 特征子空间分析方法	322
10.1 特征子空间	322
10.1.1 特征子空间的性质	322
10.1.2 子空间的比较	326
10.2 噪声子空间分析方法	328
10.2.1 Pisarenko 谐波分解	328
10.2.2 极小范数方法	330
10.3 多重信号分类 (MUSIC)	334
10.3.1 白噪声情况下的 MUSIC	334
10.3.2 有色噪声情况下的 MUSIC	336
10.4 基于修正信号子空间的波束形成器	337
10.4.1 基于特征子空间的波束形成器	338
10.4.2 基于修正信号子空间的波束形成器	339
10.4.3 波束形成器的权重向量简化	343
10.5 ESPRIT 方法	344
10.5.1 基本 ESPRIT 方法	344
10.5.2 ESPRIT 方法的拓广	347
10.5.3 广义特征值分解的 SVD-TLS 实现	349
10.6 子空间拟合法	350
10.6.1 子空间拟合问题	350
10.6.2 子空间拟合方法	351
10.7 广义相关分析	354
10.7.1 问题的描述	354
10.7.2 广义相关分解	356
10.7.3 广义 Hermitian 矩阵与特征投影算子	358
10.7.4 特征空间的渐近性质	360
10.7.5 在到达波方向估计中的应用	362
10.8 子空间约束与时频综合	365
10.8.1 子空间约束综合	366
10.8.2 诱导的自相关函数域子空间	367
10.8.3 时频综合算法	370
10.8.4 Karhunen-Loeve 展开	371
10.8.5 无基综合方法	373
参考文献	374

第十一章 子空间跟踪与更新	378
11.1 引言	378
11.2 基于 URV 分解的噪声子空间跟踪	379
11.2.1 URV 分解	380
11.2.2 平面旋转	381
11.2.3 压缩映射与细化	382
11.2.4 更新 URV 分解	384
11.3 基于秩显露 QR 分解的噪声子空间更新	386
11.3.1 更新问题	387
11.3.2 秩显露分解方法	388
11.3.3 噪声子空间的更新	391
11.4 基于一阶扰动的自适应特征值分解	393
11.4.1 秩 1 更新与扰动	394
11.4.2 一阶扰动分析	395
11.4.3 自适应特征值分解算法	396
11.5 修正特征值分解及其递推更新	401
11.5.1 修正特征值问题	402
11.5.2 秩 1 修正	406
11.5.3 秩 2 修正	407
11.6 特征子空间估计的随机梯度法	408
11.6.1 特征子空间计算的最优化理论框架	408
11.6.2 LMS 型算法	411
11.7 共轭梯度特征结构跟踪	413
11.7.1 共轭梯度法简述	413
11.7.2 代价函数	414
11.7.3 特征值分解迭代的共轭梯度算法	415
11.7.4 特征结构跟踪的共轭梯度算法	417
11.8 投影逼近子空间跟踪	420
11.8.1 信号子空间的新解释	420
11.8.2 子空间跟踪	421
11.9 快速子空间分解	424
11.9.1 Rayleigh-Ritz 逼近	424
11.9.2 基于三 Lanczos 迭代的快速子空间分解	427
11.9.3 基于双 Lanczos 迭代的快速子空间分解	428
11.10 基于 QR 分解的奇异值分解更新	430
11.10.1 奇异值分解更新	430
11.10.2 包括重新正交化的 SVD 更新	432
参考文献	433
索引	436

第一章 广义逆矩阵与 Kronecker 积

线性代数是描述许多工程问题中的数学关系所不可缺少的工具。范数、逆矩阵、特征值问题和广义逆矩阵更是信号处理和系统理论中最基本的数学工具。本章 1.1~1.5 节将简要地归纳和介绍这些常用的矩阵代数概念和结果，它们在全书中将反复、经常地用到。其中，我们将重点介绍矩阵求逆引理、广义特征值问题、广义逆矩阵的阶数递推计算以及基于广义逆矩阵的超定二维超越方程的求解。在 1.6 节，我们还将介绍在信号处理和系统理论中极为重要的一种矩阵乘积——Kronecker 积。

1.1 基本概念与符号

先给出本书中最经常使用的基本概念与符号，其中包括基本矩阵运算，独立性、子空间、基与维数，值域、零空间与秩，以及向量内积与外积。

1.1.1 矩阵符号与基本矩阵运算

令 R 表示实数集合， C 表示复数集合。一个复矩阵定义为按照长方阵列排列的复数集合，记作

$$A \in C^{m \times n} \iff A = \{a_{ij}\} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad a_{ij} \in C \quad (1.1.1)$$

式中， $C^{m \times n}$ 代表所有 $m \times n$ 复矩阵的向量空间。类似地，一个实矩阵记作

$$A \in R^{m \times n} \iff A = \{a_{ij}\}, \quad a_{ij} \in R$$

在本书中，我们经常会用到矩阵符号 $A(i_1 : i_p, j_1 : j_q)$ ，它代表由 A 的第 $i_1 \sim i_p$ 行和第 $j_1 \sim j_q$ 列组成的一个子矩阵。例如

$$A(3 : 6, 2 : 4) = \begin{bmatrix} a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \\ a_{52} & a_{53} & a_{54} \\ a_{62} & a_{63} & a_{64} \end{bmatrix}$$

特别地，当 $m = 1$ 和 $n = 1$ 时，式 (1.1.1) 分别给出复行向量

$$x \in C^{1 \times n} \iff x = [x_1, \cdots, x_n]$$

和复列向量

$$x \in C^m \iff x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, \quad x_i \in C$$

在本书中，我们将复列向量空间 $C^{m \times 1}$ 和实列向量空间 $R^{m \times 1}$ 分别简记为 C^m 和 R^m 。

分块矩阵是一个以矩阵作元素的矩阵:

$$A = \{A_{ij}\} = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$$

矩阵的基本运算包括转置、加法、标量与矩阵的乘法 $C = \alpha A$ 、矩阵与矩阵的乘法以及共轭转置 ($C^{m \times n} \rightarrow C^{n \times m}$)

$$C = A^H \implies c_{ij} = a_{ji}^*$$

其中 $*$ 代表复数共轭. 共轭转置又叫 Hermitian 伴随、Hermitian 转置或 Hermitian 共轭. 上述基本运算是矩阵计算的积木块.

除了这些矩阵基本运算外, 还有一些特殊形式的矩阵和与矩阵乘积.

具有形式为

$$A = \begin{bmatrix} A_{11} & & 0 \\ & A_{22} & \\ 0 & & A_{nn} \end{bmatrix}$$

的矩阵称为分块对角矩阵. 形式上, 这种矩阵常常用 $A = A_{11} \oplus A_{22} \oplus \cdots \oplus A_{nn}$ 表示, 或简记作 $\oplus \sum_{i=1}^n A_{ii}$, 并称 A 是矩阵 $A_{11}, A_{22}, \dots, A_{nn}$ 的直和.

一个特别简单、形式上又自然的矩阵“乘法”是按对应分量相乘. $n \times n$ 矩阵 $A = [a_{ij}]$ 和 $B = [b_{ij}]$ 的 Hadamard 积记作 $A \circ B$, 定义为 $A \circ B = [a_{ij} b_{ij}]$, 它仍然是一个 $n \times n$ 矩阵. Hadamard 积也叫 Schur 积. 像矩阵加法一样, Hadamard 积是可交换的, 因此比普通的矩阵乘法要简单得多. 有关 Hadamard 积的详细讨论可参考文献 [1] 中的 7.5 节.

另一种重要的矩阵乘积是 Kronecker 积, 我们将在 1.6 节专门讨论它.

容易证明

$$(A + B)^H = A^H + B^H$$

$$(AB)^H = B^H A^H$$

共轭转置 H 换成一般的转置 T 后, 上述二式仍然成立.

一个 $m \times n$ 分块矩阵 A 的共轭转置是一个由 A 的元素的共轭转置组成的 $n \times m$ 分块矩阵:

$$A^H = \begin{bmatrix} A_{11}^H & \cdots & A_{m1}^H \\ \vdots & & \vdots \\ A_{1n}^H & \cdots & A_{mn}^H \end{bmatrix}$$

如果矩阵 A 的元素 a_{ij} 都是参数 t 的函数, 则矩阵的导数定义为

$$\frac{dA}{dt} = \dot{A} = \begin{bmatrix} \frac{da_{11}}{dt} & \frac{da_{12}}{dt} & \cdots & \frac{da_{1n}}{dt} \\ \frac{da_{21}}{dt} & \frac{da_{22}}{dt} & \cdots & \frac{da_{2n}}{dt} \\ \vdots & \vdots & & \vdots \\ \frac{da_{m1}}{dt} & \frac{da_{m2}}{dt} & \cdots & \frac{da_{mn}}{dt} \end{bmatrix}$$

同样可定义矩阵的高阶导数.

矩阵的积分定义为

$$\int A dt = \begin{bmatrix} \int a_{11} dt & \int a_{12} dt & \cdots & \int a_{1n} dt \\ \int a_{21} dt & \int a_{22} dt & \cdots & \int a_{2n} dt \\ \vdots & \vdots & & \vdots \\ \int a_{m1} dt & \int a_{m2} dt & \cdots & \int a_{mn} dt \end{bmatrix}$$

同样也可定义矩阵的多重积分.

下面是矩阵函数的计算:

(1) 指数矩阵函数

$$e^{At} = I + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \cdots$$

(2) 指数矩阵函数的导数

$$\frac{d}{dt} e^{At} = A e^{At} = e^{At} A$$

(3) 矩阵乘积的导数

$$\frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt}$$

(4) 向量函数的导数

令 $x = [x_1, \cdots, x_n]^T$, $a = [a_1, \cdots, a_n]^T$ 和 $A \in R^{n \times n}$, 则

$$\frac{d(a^T x)}{dx} = a \quad (1.1.2)$$

和

$$\frac{d(x^T A x)}{dx} = Ax + A^T x \quad (1.1.3)$$

(5) 矩阵函数的偏导数^[9]

令 $P(x) \in R^{M \times M}$ 是复向量 $x = [x_1, \cdots, x_N]^T$ 的矩阵实函数, 且 $y = [y_1, \cdots, y_M]^T$, 则

$$\frac{\partial P(x)^{-1}}{\partial x_i} = -P(x)^{-1} \frac{\partial P(x)}{\partial x_i} P(x)^{-1} \quad (1.1.4)$$

$$\frac{\partial P(x)}{\partial x_i} y = \begin{bmatrix} \sum_{j=1}^M \frac{\partial P_{1j}(x)}{\partial x_i} y_j \\ \vdots \\ \sum_{j=1}^M \frac{\partial P_{Mj}(x)}{\partial x_i} y_j \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^M p_{1ji} y_j \\ \vdots \\ \sum_{j=1}^M p_{Mji} y_j \end{bmatrix} \quad (1.1.5)$$

和

$$\frac{\partial P(x)}{\partial x} y = \sum_{j=1}^M p_{mjn} y_j, \quad m = 1, \cdots, M; n = 1, \cdots, N \quad (1.1.6)$$

其中

$$p_{mjn} = \frac{\partial P_{mj}}{\partial x_n} \quad (1.1.7)$$

式 (1.1.2) 和式 (1.1.3) 是重要的, 因为在信号处理和系统理论的许多问题中, 往往需要使某个代价函数 (成本函数) 极小化, 而这一代价函数就包含了式 (1.1.2) 和式 (1.1.3) 的形式.

例 1.1.1 最小均方 (LMS) 自适应滤波器如图 1.1.1 所示.

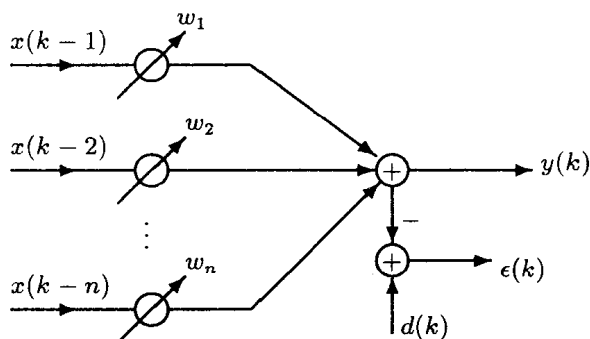


图 1.1.1 LMS 自适应滤波器

图中, 滤波器输出是输入加权后的线性组合:

$$y(k) = \sum_{i=1}^n w_i x(k-i)$$

$d(k)$ 代表“期望响应”, 输出误差定义为

$$\begin{aligned} \epsilon(k) &= d(k) - y(k) = d(k) - \sum_{i=1}^n w_i x(k-i) \\ &= d(k) - w^T x = d(k) - x^T w \end{aligned}$$

式中, $w = [w_1, \dots, w_n]^T$, 且 $x = [x(1), \dots, x(n)]^T$. 考虑均方误差

$$E\{\epsilon^2(k)\} = E\{d^2(k)\} - 2E\{d(k)x^T\}w + w^T E\{xx^T\}w$$

定义互相关函数向量 $r_{dx} = E\{d(k)x\}$ 和自相关函数矩阵 $R_{xx} = E\{xx^T\}$, 则均方误差可表述为

$$E\{\epsilon^2(k)\} = E\{d^2(k)\} - 2r_{dx}^T w + w^T R_{xx} w$$

利用式 (1.1.2) 和式 (1.1.3) 以及均方误差的最小化条件, 则有

$$\frac{\partial E\{\epsilon^2(k)\}}{\partial w} = -2r_{dx} + 2R_{xx}w = 0$$

其中, 我们使用了 R_{xx} 是对称矩阵这一事实, 即 $R_{xx}w = R_{xx}^T w$. 因此, 可求出最佳权系数向量

$$w_{\text{opt}} = R_{xx}^{-1} r_{dx}$$

这一方程式在信号处理中被称作 Wiener-Hopf 方程.

矩阵函数的偏导数在矩阵函数的极小化问题中起着关键作用, 有关式 (1.1.4)~(1.1.7) 的具体应用例子将在第十章 10.4 节介绍.

1.1.2 独立性、子空间、基与维数

域 F 上的向量空间是一些对象 (称为向量) 的集合. 向量空间 S 的子空间 S_i 是 S 的非空子集, 而该子集本身是同一个纯量域上的向量空间.

定义 1.1.1 C^m 内的向量组 $\{x_1, \dots, x_k\}$ 称为线性独立的, 若

$$\alpha_1 x_1 + \dots + \alpha_k x_k = 0$$

意味着 $\alpha_1 = \dots = \alpha_k = 0$. 反之, 若 x_i 的一个非平凡的组合等于零, 就称 $\{x_1, \dots, x_k\}$ 线性相关.

定义 1.1.2 给定向量组 $x_1, \dots, x_n \in C^m$, 则这些向量的所有线性组合的集合是一个子空间, 并称作 $\{x_1, \dots, x_n\}$ 的张成:

$$\text{span}\{x_1, \dots, x_n\} = \left\{ \sum_{j=1}^n \beta_j x_j : \beta_j \in C \right\} \quad (1.1.8)$$

在任何情况下 $\text{span}(S)$ 都是子空间, 即使 S 不是一个子空间. 若 $\text{span}(S) = V$, 则称集合张成向量空间 V . 如果 $\{x_1, \dots, x_n\}$ 是独立的, 并且 $b \in \text{span}\{x_1, \dots, x_n\}$, 那么 b 就是 x_j 的唯一线性组合.

如果 $S_1, \dots, S_k$ 是 C^m 的子空间, 那么它们的和就是由 $S = \{x_1 + x_2 + \dots + x_k : x_i \in S_i, i = 1, \dots, k\}$ 定义的子空间. 若每一个 $v \in S$ 都具有唯一的表达式 $v = x_1 + \dots + x_k$, 且 $x_i \in S_i$, 则称 S 为直和, 记作

$$S = S_1 \oplus \dots \oplus S_k$$

另外, S_i 的交集也是子空间, 记作

$$S = S_1 \cap S_2 \cap \dots \cap S_k$$

若子集 $\{x_{i_1}, \dots, x_{i_k}\}$ 是线性独立的, 且不包含在 $\{x_1, \dots, x_n\}$ 的任何其它线性独立子集内, 则称子集 $\{x_{i_1}, \dots, x_{i_k}\}$ 是 $\{x_1, \dots, x_n\}$ 的最大线性独立子集. 如果 $\{x_{i_1}, \dots, x_{i_k}\}$ 是 $\{x_1, \dots, x_n\}$ 的最大线性独立子集, 则 $\text{span}\{x_1, \dots, x_n\} = \text{span}\{x_{i_1}, \dots, x_{i_k}\}$, 并称线性独立组 $\{x_{i_1}, \dots, x_{i_k}\}$ 是 $\text{span}\{x_1, \dots, x_n\}$ 的一个基.

基不是唯一的, 但是它非常有用. 如果 $S \subseteq C^m$ 是一个子空间, 则可以找到线性独立的基向量 $x_1, \dots, x_k \in S$ 使得 $S = \text{span}\{x_1, \dots, x_k\}$.

如果向量空间 S 的某些基由有限个元素组成, 那么所有这些基将具有相同的元素个数, 该共同个数称为向量空间 S 的维数, 并记作 $\dim(S)$. 在这种情况下, 我们称 S 是有限维的; 否则, 就说 S 是无限维的.

1.1.3 值域、零空间与秩

有两类重要的子空间与一个 $m \times n$ 矩阵 A 密切相关, 它们是矩阵的值域和零空间. 矩阵 A 的值域定义为

$$\text{range } A = \{y \in C^m : y = Ax, \text{ 对某个 } x \in C^n\} \quad (1.1.9)$$