HZ BOOKS
华章教育

国外电子与电气工程技术丛书

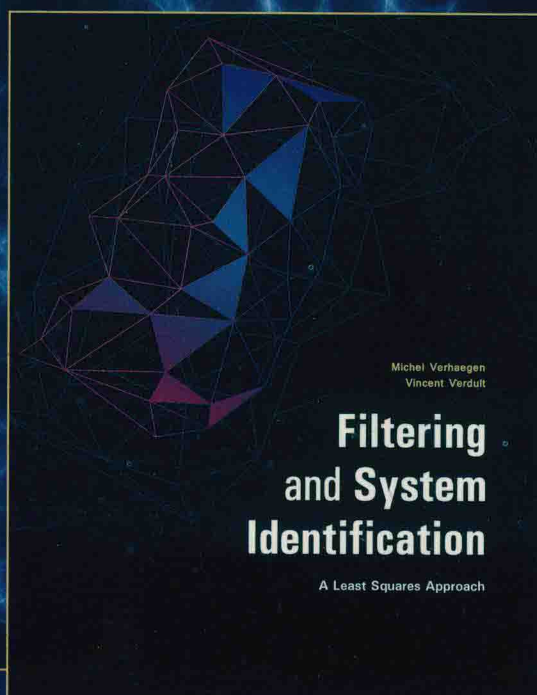
CAMBRIDGE

# 滤波与系统辨识

## 最小二乘法

[荷] 米歇尔·沃哈根（Michel Verhaegen）
文森特·沃达特（Vincent Verdult） 著
廖桂生 兰岚 廖瑞乾 刘永军 等译

*Filtering and System Identification*
*A Least Squares Approach*




机械工业出版社
China Machine Press

国外电子与电气工程技术丛书

# 滤波与系统辨识

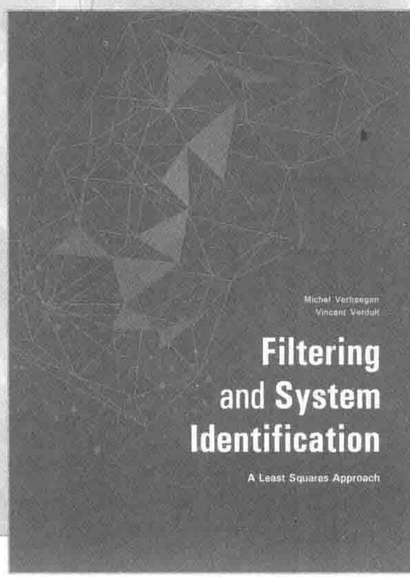
## 最小二乘法

[荷] 米歇尔·沃哈根（Michel Verhaegen）
文森特·沃达特（Vincent Verdult） 著

廖桂生 兰岚 廖瑞乾 刘永军 等译

*Filtering and System Identification*

*A Least Squares Approach*




机械工业出版社
China Machine Press

**图书在版编目（CIP）数据**

滤波与系统辨识：最小二乘法 /（荷）米歇尔·沃哈根（Michel Verhaegen），（荷）文森特·沃达特（Vincent Verdult）著；廖桂生等译．—北京：机械工业出版社，2018.7
（国外电子与电气工程技术丛书）
书名原文：Filtering and System Identification: A Least Squares Approach

ISBN 978-7-111-60647-5

I. 滤… II. ①米… ②文… ③廖… III. ①滤波技术－最小二乘法 ②系统辨识－最小二乘法 IV. ① TN713-05 ② N945.14-05

中国版本图书馆 CIP 数据核字（2018）第 181272 号

**本书版权登记号：图字 01-2016-0770**



滤波和系统识别在通信、信号处理、控制以及其他工程领域中都有很重要的应用，用于对复杂系统建模尤其重要。本书不仅讨论了如何在这些领域中设计可靠的数值方法来重构系统的未知信息，还讨论了在线性状态－空间模型中如何利用最小二乘法提升估计方法日益复杂的问题。本书适合电子、机械等专业的研究生和研究人员阅读，也可供相关专业人士参考。

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）

| | | | |
|---|---|---|---|
| 责任编辑： | 谢晓芳 | 责任校对： | 李秋荣 |
| 印　　刷： | 北京文昌阁彩色印刷有限责任公司 | 版　　次： | 2018 年 9 月第 1 版第 1 次印刷 |
| 开　　本： | 185mm × 260mm　1/16 | 印　　张： | 13.25 |
| 书　　号： | ISBN 978-7-111-60647-5 | 定　　价： | 69.00 元 |

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991　88361066　　投稿热线：（010）88379604
购书热线：（010）68326294　88379649　68995259　　读者信箱：hzjsj@hzbook.com

# 出版者的话

文艺复兴以来，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，信息学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的信息产业发展迅猛，对专业人才的需求日益迫切。这对我国教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其信息科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀教材将对我国教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到"出版要为教育服务"。自 1998 年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与 Pearson、McGraw-Hill、Elsevier、John Wiley & Sons、CRC、Springer 等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出 Alan V. Oppenheim、Thomas L. Floyd、Charles K. Alexander、Behzad Razavi、John G. Proakis、Stephen Brown、Allan R. Hambley、Albert Malvino、Peter Wilson、H. Vincent Poor、Hassan K. Khalil、Gene F. Franklin、Rex Miller 等大师名家的经典教材，以"国外电子与电气工程技术丛书"和"国外工业控制与智能制造丛书"为系列出版，供读者学习、研究及珍藏。这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版"经典原版书库"作为姊妹篇也越来越多被实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着电气与电子信息学科、自动化、人工智能等建设的不断完善和教材改革的逐渐深化，教育界对国外电气与电子信息类、控制类、智能制造类等相关教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方法如下：


华章网站：www.hzbook.com

电子邮件：hzjsj@hzbook.com

联系电话：(010)88379604

联系地址：北京市西城区百万庄南街 1 号

邮政编码：100037


HZ BOOKS
华章教育

华章科技图书出版中心

# 译 者 序

在通信、信号处理、控制和其他工程学科中，滤波和系统识别为建立复杂系统模型提供了强有力的技术支撑。本书讨论了在利用这些技术导出的模型中，如何设计可靠的数学方法来获取未知变量，尤其集中在线性状态空间模型中利用最小二乘法来获取未知变量的渐增的复杂估计问题。

本书作者从线性矩阵代数、信号变换、线性系统理论和随机变量等关键的基础问题入手，包括在状态空间模型中的多种估计和识别方法。从卡尔曼滤波器开始，对大量滤波和系统识别问题进行了分析，直接从数据得到一个完整的模型、噪声统计特性以及状态估计。最后一章关于系统识别循环为读者解决实际问题做了详细阐述。

本书章节后的习题、MATLAB 仿真以及图例适用于电气、机械和航空航天工程领域的研究生以及科研人员，同时也为从业者提供了有用参考。关于此内容的额外资料，包括指导方案，可在网站 www.cambridge.org/9780521875127 上查到。

本书作者 Michel Verhaegen 是荷兰代尔夫特理工大学教授、代尔夫特系统和控制中心(DCSC)主任。他目前的研究包括应用于工业的基准的新辨识和控制器设计的方法，特别侧重于自适应光学、主动振动控制和底盘全方位控制系统等领域。

本书作者 Vincent Verdult 于 2001—2005 年在荷兰代尔夫特理工大学任助理教授，主要研究非线性状态空间系统的系统辨识。他目前的研究方向是信息论。

本书译者都是多年工作在信号处理技术领域的一线研究人员，但是由于滤波与系统识别涉及的知识范围广，所以我们对于原著内容的理解难免会存在偏差，翻译不当之处，希望得到各位同行和专家的批评指正。

译 者

2018 年 6 月

# 前　言

本书主要阐述了线性代数和最小二乘问题在滤波和系统辨识领域的作用，适用于工程领域的研究生一年级的学生。从荷兰代尔夫特理工大学和特温特大学获得的经验来看，重读本科时期的线性代数、统计学和系统论等课程的相关资料，即可让学习本书成为进入研究生课程的理想开端。更为重要的是，线性代数中的几何概念和最小二乘的重要作用可以鼓舞学生们去了解滤波和系统辨识算法是怎样引出的，进而启蒙他们寻求新的算法。同时，本书也给学生们提供了一个了解应用数学是如何解决实际工程问题的机会。

本书内容可分为七讲

(1) 第一讲：线性代数的介绍与回顾(第 1、2 章)

(2) 第二讲：系统论和概率论回顾(第 3、4 章)

(3) 第三讲：卡尔曼滤波(第 5 章)

(4) 第四讲：频率响应函数估计(第 6 章)

(5) 第五讲：状态空间的参数估计(第 7、8 章)

(6) 第六讲：子空间模型辨识(第 9 章)

(7) 第七讲：从理论到实际应用：系统辨识环(第 10 章)

本书作者认为，每讲之后，还应开设习题课，课上学生们在导师的指导下做一些相关的练习，这样传授知识的效果能显著地提高。同时，在习题课上，每个学生都有机会就与本课程相关的问题进行提问。在荷兰代尔夫特理工大学，本课程的教授方式是一个真实的案例研究，即应用本书涵盖的内容，通过测量输入和输出数据来确认一个数学模型。

本书作者已经在荷兰代尔夫特理工大学和特温特大学使用了这本书给硕士生讲课。来听课的学生来自各个系，包括：电子、机械、航天工程和应用物理。如今，作为滤波和辨识的基础课，本书已经成为代尔夫特大学系统与控制中心在系统与控制方面硕士课程体系的重要组成部分(http://www.dcsc.tudelft.nl)。本书的一部分也用于代尔夫特系统与控制研究所(DISC)的研究生教育项目。另外，Bernard Hanzon 在奥地利维也纳技术大学担任访问学者时用了本书的一部分，Jonas Sjöberg 在瑞典查尔姆斯特理工大学时也把本书的一部分内容用于本科教学。

作者写这本书的缘由是希望他们的学生能像他们自己一样对滤波和系统辨识领域充满热情。尽管在书稿创作过程中这些学生起到了促进和核心作用，但业界同事通过紧密合作才实现了本书的最终版式和质量。为此，作者们对以下人士的建设性意见表示感谢，他们是：Dietmar Bauer (Technische Universtät Wien, Austria), Bernard Hanzon (University College Cork, Ireland), Gjerrit Meinsma (University of Twente, the Netherlands), Petko Petkov (Technical University of Sofia, Bulgaria), Philip Regalia (Institut National des Télécommunications , France), Ali Sayed (University of California, Los Angeles, USA), Johan Schoukens (Free University of Brussels, Belgium), Jonas Sjöberg (Chalmers University of Technology, Sweden), and Rufus Fraanje (TU Delft).

特别感谢 Niek Bergboer (荷兰马斯特里赫特大学)在开发 Matlab 软件和使用指南中做出的巨大贡献，非常有助于对本书辨识方法的描述。最后感谢博士生 Paolo Massioni 和 Justin Rice 在校对和本书习题答案方面做出的努力。

关于本书教辅资源，只有使用本书作为教材的教师才可以申请，需要的教师可向剑桥大学出版社北京代表处申请，电子邮件 solutions@cambridge.org。

# 符号和表示

| | |
|---|---|
| $\mathbb{Z}$ | 整数集 |
| $\mathbb{N}$ | 正整数集 |
| $\mathbb{C}$ | 复数域 |
| $\mathbb{R}$ | 实数域 |
| $\mathbb{R}^n$ | 实 $n$ 维向量空间 |
| $\mathbb{R}^{m\times n}$ | 实 $m\times n$ 维矩阵空间 |
| $\infty$ | 无穷 |
| Re | 实部 |
| Im | 虚部 |
| $\in$ | 属于 |
| $=$ | 等于 |
| $\approx$ | 约等于 |
| $\square$ | 证毕 |
| $\otimes$ | 克罗内克积 |
| $I_n$ | $n\times n$ 单位矩阵 |
| $[A]_{i,j}$ | 矩阵 $A$ 的第 $i$ 行第 $j$ 列元素 |
| $A(i,:)$ | 矩阵 $A$ 的第 $i$ 行 |
| $A(:,i)$ | 矩阵 $A$ 的第 $i$ 列 |
| $A^{\mathrm{T}}$ | 矩阵 $A$ 的转置 |
| $A^{-1}$ | 矩阵 $A$ 的逆 |
| $A^{1/2}$ | 矩阵 $A$ 的对称正定平方根 |
| $\mathrm{diag}(a_1, a_2, \cdots, a_n)$ | 对角元素为 $a_i$ 的 $n\times n$ 单位矩阵 |
| $\det(A)$ | 矩阵 $A$ 的行列式 |
| $\mathrm{range}(A)$ | 矩阵 $A$ 的列空间 |
| $\mathrm{rank}(A)$ | 矩阵 $A$ 的秩 |
| $\mathrm{trace}(A)$ | 矩阵 $A$ 的迹 |
| $\mathrm{vec}(A)$ | 矩阵 $A$ 按行拉直所得的列向量 |
| $\\|A\\|_2$ | 矩阵 $A$ 的 2 范数 |
| $\\|A\\|_{\mathrm{F}}$ | 矩阵 $A$ 的弗罗贝尼乌斯范数 |
| $\\|x\\|_i$ | 向量 $x$ 的第 $i$ 个元素 |
| $\\|x\\|_2$ | 向量 $x$ 的 2 范数 |
| lim | 极限 |
| min | 最小 |
| max | 最大 |
| sup | 最小上界 |
| $\mathrm{E}[\cdot]$ | 统计期望值 |
| $\delta(t)$ | 冲激函数 |
| $\Delta(k)$ | 单位脉冲函数 |
| $\mathrm{s}(k)$ | 单位阶跃函数 |
| $X\sim(m,\sigma^2)$ | 均值为 $m$、方差为 $\sigma^2$ 的高斯随机变量 |

# 缩略语对照表

| | | |
|---|---|---|
| ARX | Auto-Regressive with eXogeneous input | 外源输入自回归 |
| ARMAX | Auto-Regressive Moving Average with eXogeneous input | 外源输入自回归滑动平均 |
| BIBO | Bounded Input, Bounded Output | 有界输入，有界输出 |
| BJ | Box-Jenkins | 人名 |
| CDF | Cumulative Distribution Function | 累积分布函数 |
| DARE | Discrete Algebraic Ricatti Equation | 离散代数里卡蒂方程 |
| DFT | Discrete Fourier Transform | 离散傅里叶变换 |
| DTFT | Discrete-Time Fourier Transform | 离散时间傅里叶变换 |
| ETFE | Empirical Transfer-Function Estimate | 经验传递函数估计 |
| FFT | Fast Fourier Tansform | 快速傅里叶变换 |
| FIR | Finite Impulse Response | 有限冲激响应 |
| FRF | Frequency-Response Function | 频率响应函数 |
| IID | Independent, Identically Distributed | 独立同分布 |
| IIR | Infinite Impulse Response | 无限冲激响应 |
| LTI | Linear Time-Invariant | 线性时不变 |
| LTV | Linear Time-Varying | 线性时变 |
| MIMO | Multiple Input, Multiple Output | 多输入多输出 |
| MOESP | Multivariable Output-Error State-Space | 多变量输出误差状态空间 |
| N4SID | Numerical algorithm for Subspace IDentification | 子空间辨识的数值算法 |
| PDF | Probability Density Function | 概率密度函数 |
| PEM | Prediction-Error Method | 预测误差法 |
| PI | Past Inputs | 过去输入 |
| PO | Past Outputs | 过去输出 |
| OE | Output-Error | 输出误差 |
| RMS | Root Mean Square | 方均根 |
| SISO | Single Input, Single Output | 单输入单输出 |
| SRCF | Square-Root Covariance Filter | 协方差平方根滤波器 |
| SVD | Singular-Value Decomposition | 奇异值分解 |
| WSS | Wide-Sense Stationary | 广义平稳性 |

# 目　录

# 第1章 概　论

通过对我们周围环境的感觉来进行观察是生物的自然行为。生物获取的信息也是多样、连续的，例如声音信号和图像信号。这些信息被处理并用来产生一个适用于当前状况的特定环境模型。这种基于观察而建模的行为深植于我们人类天性中，并影响着我们日常生活中所做的决定和选择。

通过观察来建模的方法在科学的许多分支中也起着重要的作用。虽然利用我们的感觉在观察中的作用举足轻重，然而科学观察通常是利用测量仪器和传感器实现的。这些传感器获得的测量数据通常要经过处理才能用来评判或验证这项观测实验，或者为了获得更多的信息来引导实验。获取的数据通常用于建立一个能描述实验动态特性的数学模型。系统辨识方法是由测量数据来建立数学模型的一类有体系的方法。得到的数学模型的一个重要应用就是通过对获得的测量结果做滤波处理来预测模型的数量。

在滤波和系统辨识领域中，1800 年前 Johann Carl Friedrich Gauss(1777—1855)提出的最小二乘法是一个历史的里程碑。在本书中，最小二乘法运用在滤波和系统辨识中是一个反复出现的主题。本书也给出了最小二乘法的早期发展的历史进程描述简介，该描述参照 Bühler 给出的概要(1981)。

高斯(Gauss)当初提出最小二乘法的时候，并没有认识到它的重要性。第一份关于最小二乘法的出版物还是 Adrien Marie Legendre(1752—1833)在 1806 年所著的，而在那之前，高斯已经非常明确和频繁地使用这方法了。高斯为此受到激发，在论文“Theoria combinationis observationum erroribus minimis obnoxiae Ⅰ和Ⅱ”(1821 和 1823)中翔实地推导了最小二乘法。论文Ⅰ致力于介绍该理论，论文Ⅱ则包含了实际应用，其中大部分是关于航天问题的。在论文Ⅰ中，高斯提出了一个关于偶然误差(Zufallsfehler)概率理论，他定义了观测 $x$ 的误差概率分布函数 $\phi(x)$，$\phi(x)\mathrm{d}x$ 是误差在 $x$ 到 $x+\mathrm{d}x$ 区间之间的概率。函数 $\phi(x)$满足标准化条件

$$\int_{-\infty}^{+\infty}\phi(x)\mathrm{d}x = 1$$

高斯给出的一个决定性的要求就是下面积分要取到最小值

$$\int_{-\infty}^{+\infty}x^2\phi(x)\mathrm{d}x$$

选择当误差的平方取到最小值时作为最合适的权，这也就是为什么这个方法被称作最小二乘(最小平方)。但是，这种选择的方法遭到了皮尔·西蒙·拉普拉斯(Pierre Simon Laplace，1749—1827)的质疑，他早期使用的是误差的绝对值，然而平方的方法从计算方面讲显然优于拉普拉斯(Laplace)最初的方法。

在提出最小二乘的基础理论之后，高斯就要找出一个合适的函数 $\phi(x)$。这时，高斯引进了高斯分布从而给出了观测误差“自然”分布的形式。

$$\phi(x) = \frac{1}{\pi}\mathrm{e}^{-x^2}$$

高斯未曾在论文中提到统计分布函数与高斯分布的不同。他满足于自己的成就，对自己的理论的应用并没有激励他去寻找其他分布函数。最小二乘法在 19 世纪初是高斯解决实验研究问题中不可缺少的理论工具，他自己也认为这是数学和自然之间联系的最重要的见证。

直至今日，最小二乘法在数学建模里的影响也是巨大的，并且在任何有关这个主题的书中，都只是归结到一些限定性问题中。在本课程中，我们主要关注的是通过测量工程系统的输入和输出的数据序列，来识别线性状态空间模型。这么讲可能起初看起来并没有我们想要建模的囊括系统辨识的主要组成部分，其实不然。随后，作为特例，本书将阐释状态空间方法是可以通过估计函数中的参数来解决现有的识别问题的，并以广泛应用的ARX和ARMAX模型(Ljung，1999)来做例子。

本书的主要目标是帮助读者发现线性最小二乘法是如何解决线性状态空间模型识别问题的。线性最小二乘可以表示为如下的确定性参数优化问题：

$$\min_{x} \mu^{\mathrm{T}}\mu \text{ 满足 } y = Fx + \mu \tag{1.1}$$

其中：向量 $y \in \mathbb{R}^{N}$；已知矩阵 $F \in \mathbb{R}^{N\times n}$；$x \in \mathbb{R}^{n}$ 是需要确定的未知向量。这个优化问题的解在很多教材中都可以找到。虽然用几行式子就能找到上式的解析解，但大量的教材却从不同的角度分析了如何求解。很多例子都是在大量的假设下利用统计知识阐释这个解析解的，利用矩阵 $F$ 的元素和向量 $\mu$ 或者利用矩阵 $F$ 结构时得到的是数字解。在更深入的最小二乘问题和其应用中，我们参考了 Kailath 等人的书(2000)。

在开始主要课程之前，本书前三章主要介绍基础。第2章给出了较新的矩阵线性代数研究。第3章给出了确定信号系统的信号传输方式及其线性系统理论。第4章主要针对随机变量和随机信号。要想理解本书提到的系统辨识方法，必须掌握前三章背景内容的深刻内涵。

通常，识别一个动态系统第一步就是确定一个预测器。所以，在第5章，我们首先研究了如何去预测一个线性状态空间模型。除了要知道输入和输出，在做状态预测或状态观测时还需要知道动态模型的相关知识(状态空间形式)，随机扰动的均值和协方差矩阵。优化状态重建问题可以定义为一个最小二乘的场景。在第5章，最优预测或卡尔曼(Kalman)滤波问题可以表示和解答为一个加权线性最小二乘问题。这个表达和解答由Paige在1985年首先提出。这种表示方式最大的优点就是把一个递归问题从基础性的线性代数概念推导出来，例如用高斯消元法解一组超定方程。我们还会简要地讨论卡尔曼滤波在估计动态系统的未知输入问题上的应用。

第6章讨论了在频域估计线性状态模型的输入-输出的描述问题。估计这些描述，比如说频率响应函数(FRF)，都基于时间序列的(离散)傅里叶(Fourier)变换。这一章的讨论包括有限持续的实验的实际约束对FRF估计精度的影响。同样简要介绍了快速傅里叶变换(FFT)在推导快速算法方面的应用。频域方法在处理大量数据时的一大优势是可以使用快速算法。在工业的主要组成部分，例如汽车和飞机制造业，该方法是重获动态系统信息的主要工具。

第7章讨论了假设输出观测在加性白噪声背景下，且状态向量有固定已知的维度时，状态空间模型的系统矩阵元素的估计问题。紧接着，引出了所谓的输出误差方法(Ljung，1999)。这个基础性的估计问题揭示了一系列辨识问题的核心和关键。解决这个问题最先要解决的是如何把系统矩阵的元素表示为一个个参数未知的向量函数。本书把如何选择这个参数向量看作参数化问题，建立参数化多变量状态空间模型的方法也有很多。参数一旦确定，输出误差问题就可以用如下的最小二乘问题表示：

$$\min_{x_1, x_2} \mu^{\mathrm{T}}\mu \text{ 满足 } y = F(x_1)x_2 + \mu \tag{1.2}$$

其中：$x_1$ 和 $x_2$ 是待测未知参数向量。这种类型的最小二乘问题要比式(1.1)难得多，因为矩阵 $F$ 是由未知参数 $x_1$ 决定的。通常用迭代的方法求解，因此需要给出 $x_1$ 和 $x_2$ 的初值。更进一步的，通常不能保证这样的数字迭代步骤收敛于全局最优值 $\mu^{\mathrm{T}}\mu$。在本章，我们主要关注优化输出误差迭代步骤的数值实现方法。获得未知向量的估计之后评价估计准确性的问题是通过在无偏估计的假设下计算这些估计的协方差矩阵实现的。在本章结束时我们

讨论了如何在加性噪声非白噪声的情况下避免有偏估计的问题。

第 8 章介绍了在固定的已知状态维度的情况下来辨识预测器(卡尔曼滤波器)的经典预测误差方法(PEM)(Ljung，1999)。这个问题最终归结到卡尔曼滤波器预测模型的参数估计问题。第 7 章给出了适用于这些预测器模型的输出误差问题及求解。除了提到一般的多变量状态空间模型的预测误差法外，这里还特别提到了单输入单输出(SISO)系统。安排这些内容的目的，第一是为了说明一些广为人知的模型结构例如 ARMAX 模型可以看作是状态空间模型的一例权威的参数化。第二是说明当识别状态维度不同的系统时，可以对偏差进行量化分析。

第 9 章讲了最近提出的子空间辨识方法。这些方法可以实现多变量状态空间模型的准确估计，在一般的噪声背景下，仅需解出形式如式(1.1)那样的线性最小二乘问题即可。子空间方法无论在学术上还是工业上之所以有价值，一部分是因为在估计模型和它的阶数时并没有要求模型必须参数化。这是通过把由模型矩阵定义的子空间和由现有的观测建立起的结构矩阵联系起来而实现的。子空间在这些方法中所起到的重要作用正是这些方法被称作子空间方法的原因。一种区分这些子空间算法的方式是，看这些方法是如何使用辅助变量的概念与各种噪声场景做匹配的。虽然在实际应用中，子空间方法已经被证实可以即刻得到精确的模型，但是它们却不能像预测误差方法那样优化预测误差准则。为了实现这个统计优化，可以采用子空间方法得到估计值作为预测误差的初始值。这个概念已经被 Ljung(Math Works，2000a)提出。

第 10 章介绍了模型估计算法和它们在一个辨识实验中的应用。为了展开分析并改进实验，这里讨论了一个由 Ljung(1999)提出的循环方法。这个循环方法的目的在于对在系统辨识里的系统解决方案做出选择。这些选择包括实验条件的选择(例如采样频率、实验持续时间或者输入信号的类型)，记录的时间序列的处理(消除趋势、移除异常值和滤波)，在参数估计算法中的模型结构(选择的模型的阶数和延时)。本章还讨论了如何将第 9 章介绍的子空间方法和第 7、8 章介绍的参量方法结合起来，帮助辨识系统做出关于模型结构的选择，这也使得对多变量系统的全面识别变得可行。单独使用预测误差方法时，找到合适的模型结构需要对大量的概率数据进行测验，这实际上是不可实现的，因为不仅仅需要识别一个模型，还需要识别一系列的不同实验条件下的模型。

每一章的末尾，我们都布置了习题帮助读者实践这些算法。作者们也开发了一个包含本书提到辨识方法的 Matlab 工具包以及一个综合的软件指南(Verhaegen 等，2003)。

滤波和系统辨识是学科交叉的极好例子，它不仅在许多领域得到广泛应用，也把(数学)原理定律中很多基础性的知识结合起来了。作者认为本书现有的结构和内容只是系统辨识领域的入门。学习掌握这本教材固然好，但更重要的是能够利用本书的知识解决实际问题。

# 第2章 线性代数

在学习本章后，你将能够：

- 进行向量和矩阵的基本运算；
- 定义向量空间；
- 定义向量空间的子空间；
- 计算矩阵的秩；
- 列出由线性变换定义的四个基本子空间；
- 计算方阵的逆，行列式，特征值和特征向量；
- 描述正定矩阵；
- 计算一些重要的矩阵分解，如矩阵的特征值分解，奇异值分解和 $QR$ 因式分解；
- 利用线性代数方法求解线性方程；
- 描述确定性最小二乘问题；
- 用数值计算确定性最小二乘问题。

## 2.1 简介

在本章中，我们从线性代数中回顾一些基本问题。所讨论的内容在后续章节中会经常用到。

自 20 世纪 60 年代以来，线性代数作为技术进步的重要贡献因素，在工程中已经发挥了重要作用。线性代数是一种用数字方式解决系统理论问题(如滤波和控制问题)的工具。线性代数在工程中的广泛应用促进了该领域的发展，特别是算法的数值分析。Matlab(MathWorks，2000b)和 SciLab(Gomez，1999)等计算机辅助设计软件包的广泛使用，提升了线性代数在工程中的应用。这些软件包的用户友好性使我们能够用几行代码为复杂的系统理论问题编写解决方案。因此，这大大加快了新算法的原型设计。然而，另一方面需要注意的是，Matlab 中的编码可能给用户这样的印象：成功运行 Matlab 程序等同于充分证明一个新理论。为了避免养成“经过 Matlab 证明”的看法，本章和后续章节中线性代数中的算法主要涉及推导，而不是应用。

本章开始，我们回顾了线性代数的两个基本概念：向量和矩阵。向量在 2.2 节中介绍，矩阵在 2.3 节中介绍。对于一类特殊的矩阵，存在几个重要的概念，这些在 2.4 节中有介绍。2.5 节介绍了一些已经被证明在滤波和估计上有用的矩阵分解。最后，2.6 节和 2.7 节重点讨论最小二乘问题，其中需要解决一个超定的线性方程组。这些问题是特别有意义的，因为大量的滤波、估计，甚至与控制有关问题都可以写成线性(加权)最小二乘问题。

## 2.2 向量

向量是实数或复数的数组。在这本书中我们使用 $\mathbb{R}$ 表示实数集合，$\mathbb{C}$ 表示复数集合。向量有两种形式，列向量和行向量。由各元素组成的列向量 $x_1$，$x_2$，…，$x_n$($x_i \in \mathbb{C}$)可以表示成：

$$x=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$$

在本书中，用小写字母表示一个列向量。行向量由 $x^{\mathrm{T}}$ 表示，也就是说，向量由如下元素组成：

$$x^{\mathrm{T}}=[x_1\quad x_2\quad \cdots\quad x_n]$$

行向量 $x^{\mathrm{T}}$ 也称为列向量 $x$ 的转置。向量中元素的数量称为向量的维数。将具有 $n$ 个元素的向量称为 $n$ 维向量。我们使用符号 $x\in\mathbb{C}^n$ 表示一个 $n$ 维复向量。显然，一个 $n$ 维实向量由 $x\in\mathbb{R}^n$ 表示。在本书中，大多数向量是实值；因此，在本章的剩余部分我们将讨论范围限制在实值向量中。但是，大多数的结果可以容易扩展到复向量。

向量 $x\in\mathbb{R}^n$ 与标量 $\alpha\in\mathbb{R}$ 的乘积定义为

$$\alpha x=\begin{bmatrix}\alpha x_1\\\alpha x_2\\\vdots\\\alpha x_n\end{bmatrix}$$

两个向量 $x$，$y\in\mathbb{R}^n$ 的和定义为

$$x+y=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}+\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}=\begin{bmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{bmatrix}$$

两个向量 $x$，$y\in\mathbb{R}^n$ 的标准内积定义为

$$x^{\mathrm{T}}y=x_1y_1+x_2y_2+\cdots+x_ny_n$$

向量 $x$ 的二范数用 $\|x\|_2$ 表示，它是向量关于自身内积的平方根，即

$$\|x\|_2=\sqrt{x^{\mathrm{T}}x}$$

如果两个向量内积等于零，则称两个向量正交；如果两个向量正交且它们的二范数都为 1，则称两个向量为标准正交。

任何两个向量 $x$，$y\in\mathbb{R}^n$ 满足柯西-施瓦茨不等式：

$$|x^{\mathrm{T}}y|\leqslant\|x\|_2\,\|y\|_2$$

等式当且仅当 $x=\alpha y$，$\alpha\in\mathbb{R}$ 或者 $y=0$。

如果这些向量的任意的线性组合

$$\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_mx_m \tag{2.1}$$

($\alpha_i\in\mathbb{R}$，$i=1$，2，…，$m$)等于 0 仅当对于所有 $\alpha_i=0$，$i=1$，2，…，$m$ 成立，则称这 $m$ 个向量 $x_i\in\mathbb{R}^n$，$i=1$，2，…，$m$ 是线性独立的。如果式(2.1)等于 0 时系数 $\alpha_i$ 不全为 0，则称这 $m$ 个向量是线性相关的或共线性的。在后一种情况下，至少其中一个向量如 $x_m$ 可以由其他向量的线性组合表示，即

$$x_m=\beta_1x_1+\beta_2x_2+\cdots+\beta_{m-1}x_{m-1}$$

其中，$\beta_i\in\mathbb{R}$，$i=1$，2，…，$m-1$。注意，如果$m>n$，则 $m$ 个向量 $x_m\in\mathbb{R}^n$ 必然是线性相关的。让我们用一个例子来说明线性独立。

**例 2.1** **(线性独立)** 考虑向量

$$x_1=\begin{bmatrix}1\\1\\1\end{bmatrix},\quad x_2=\begin{bmatrix}0\\1\\0\end{bmatrix},\quad x_3=\begin{bmatrix}2\\3\\2\end{bmatrix}$$

向量 $x_1$ 和 $x_2$ 是线性独立的，因为

$$\alpha_1 x_1 + \alpha_2 x_2 = 0$$

解方程得到 $\alpha_1=0$，$\alpha_1+\alpha_2=0$，所以有 $\alpha_1=\alpha_2=0$。向量 $x_1$、$x_2$ 和 $x_3$ 是线性相关的，因为 $x_3=2x_1+x_2$。◀

一个向量空间 $\mathcal{V}$ 是一组向量的集合，它具有加法和与实数的乘法规则，满足如下性质：

(i) 对于任意 $x$，$y\in\mathcal{V}$，有 $x+y\in\mathcal{V}$；

(ii) 对于任意 $\alpha\in\mathbb{R}$，$x\in\mathcal{V}$，有 $\alpha x\in\mathcal{V}$；

(iii) 对于任意 $x$，$y\in\mathcal{V}$，有 $x+y=y+x$；

(iv) 对于任意 $x$，$y$，$z\in\mathcal{V}$，有 $x+(y+z)=(x+y)+z$；

(v) 对于任意 $x\in\mathcal{V}$，有 $x+0=x$；

(vi) 对于任意 $x\in\mathcal{V}$，有 $x+(-x)=0$；

(vii) 对于任意 $x\in\mathcal{V}$，有 $1\times x=x$；

(viii) 对于任意 $\alpha_1$，$\alpha_2\in\mathbb{R}$，$x\in\mathcal{V}$，有 $(\alpha_1\alpha_2)x=\alpha_1(\alpha_2 x)$；

(ix) 对于任意 $\alpha\in\mathbb{R}$，$x$，$y\in\mathcal{V}$，有 $\alpha(x+y)=\alpha x+\alpha y$；

(x) 对于任意 $\alpha_1$，$\alpha_2\in\mathbb{R}$，$x\in\mathcal{V}$，有 $(\alpha_1+\alpha_2)x=\alpha_1 x+\alpha_2 x$。

例如，所有 $n$ 维向量的集合形成一个向量空间。如果向量空间中的每个向量可以由一组向量 $x_i\in\mathcal{V}$，$i=1$，2，…，$\ell$ 线性表示。这些向量 $x_i$ 张成空间 $\mathcal{V}$。也就是说，空间 $\mathcal{V}$ 中的任意向量 $y\in\mathcal{V}$ 可以展开写成：

$$y = \sum_{i=1}^{\ell} \alpha_i x_i$$

其中：$\alpha_i\in\mathbb{R}$，$i=1$，2，…，$\ell$。向量空间的一组基是指其可以张成向量空间 $\mathcal{V}$ 并且它们两两线性无关。一组正交基是指基中的每个向量与所有其他向量都正交。基中的向量个数称为空间的维数。所以，一个 $n$ 维向量空间至少需要 $n$ 个向量张成。

一个向量空间 $\mathcal{V}$ 的子空间 $\mathcal{U}$，用 $\mathcal{U}\subset\mathcal{V}$ 表示，当它是非空子集时需要满足下面两条性质：

(i) 对于任意 $x$，$y\in\mathcal{U}$，有 $x+y\in\mathcal{U}$；

(ii) 对于任意 $\alpha\in\mathbb{R}$，$x\in\mathcal{U}$，有 $\alpha x\in\mathcal{U}$。

也就是说，两个向量的和或者一个标量与向量的积所生成的新的向量仍然在该向量子空间。对第二条性质中的标量取 $\alpha=0$ 可以得到零向量，它属于每个子空间。因为向量的加法和与标量乘法规则继承自向量空间 $\mathcal{V}$，每个向量子空间本身就是一个向量空间，如果两个属于相同空间 $\mathcal{V}$ 的子空间 $\mathcal{U}$ 和 $\mathcal{W}$ 中的每个向量彼此正交，则称这两个子空间正交。对于一个空间 $\mathcal{V}$ 中的子空间 $\mathcal{U}$，空间 $\mathcal{V}$ 中的所有与子空间 $\mathcal{U}$ 正交的向量构成子空间 $\mathcal{U}$ 的补空间，记做 $\mathcal{U}^{\perp}$。

**例 2.2** **(向量空间)** 向量组

$$x_1 = \begin{bmatrix}0\\1\\0\end{bmatrix},\quad x_2 = \begin{bmatrix}2\\0\\0\end{bmatrix},\quad x_3 = \begin{bmatrix}-1\\0\\-1\end{bmatrix}$$

是线性无关的，并组成三维向量空间的一组基。它们不是一组正交基，由于 $x_2$ 与 $x_3$ 不正交。$x_1$ 张成三维空间中的一维子空间即一条线。这个子空间与三维空间中的 $x_2$ 和 $x_3$ 张成的两维子空间正交，因为 $x_1$ 与 $x_2$ 和 $x_3$ 都正交。◀

## 2.3 矩阵

一个向量是具有一列或一行的数组。一个具有 $m$ 行 $n$ 列的数组称为 $m$ 乘以 $n$ 维矩阵，或者简单地说，是一个 $m\times n$ 矩阵。一个 $m\times n$ 矩阵写为：

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

其中：标量 $a_{ij}$ 称为矩阵的第$(i,\ j)$个元素。在本书中，矩阵用大写字母表示。我们使用符号 $A\in\mathbb{C}^{m\times n}$ 表示矩阵元素是复数的 $m\times n$ 矩阵，$A\in\mathbb{R}^{m\times n}$ 表示矩阵元素是实数的 $m\times n$ 矩阵。我们遇到大多数情况矩阵都是实数矩阵，所以在本章中我们将主要讨论实数矩阵。

显然，$n$ 维列向量可以看作 $n\times 1$ 维矩阵，$n$ 维行向量可以看作 $1\times n$ 维矩阵。

具有行列相同数量的矩阵称为方阵。方阵具有一些特殊的性质，我们将在 2.4 节介绍。

矩阵 $A\in\mathbb{R}^{m\times n}$ 可以看作 $n$ 列 $m$ 维列向量的集合，即

$$A=\left[\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} \cdots \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}\right]$$

或者可以看作 $m$ 行 $n$ 维行向量的集合，即

$$A=\begin{bmatrix} [a_{11} & a_{12} & a_{13} & \cdots & a_{1n}] \\ [a_{21} & a_{22} & a_{23} & \cdots & a_{2n}] \\ & & & \vdots & \\ [a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}] \end{bmatrix}$$

这表明我们可以将矩阵分割成列向量或行向量。更一般地，矩阵可划分成几个子矩阵，例如

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\in\mathbb{R}^{m\times n}$$

其中：对于 $p<m$，$q<n$，有矩阵 $A_{11}\in\mathbb{R}^{p\times q}$，$A_{12}\in\mathbb{R}^{p\times(n-q)}$，$A_{21}\in\mathbb{R}^{(m-p)\times q}$，以及 $A_{22}\in\mathbb{R}^{(m-p)\times(n-q)}$。

矩阵的转置记作 $A^{\mathrm{T}}$，是一个 $n\times m$ 维矩阵，它由矩阵 $A$ 的行列互换得到，即

$$A^{\mathrm{T}}=\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

显然，$(A^{\mathrm{T}})^{\mathrm{T}}=A$。

单位矩阵是一种特殊矩阵，记作 $I$。它是一种对角线元素非零的方阵，并且对角线元素全等于 1，即

$$I=\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{2.2}$$

如果将矩阵 $A\in\mathbb{R}^{m\times n}$ 乘以标量 $\alpha\in\mathbb{R}$，得到一个第$(i,\ j)$个元素为 $\alpha a_{ij}$ $(i=1,\ \cdots,\ m,\ j=1,\ \cdots,\ n)$的矩阵。两个相同维数的矩阵 $A\in\mathbb{R}^{m\times n}$，$B\in\mathbb{R}^{m\times n}$ 的和等于各自行列上元素之和 $a_{ij}+b_{ij}$。显然有$(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}}$。两个矩阵 $A\in\mathbb{R}^{m\times n}$，$B\in\mathbb{R}^{n\times p}$ 的乘积得到的矩阵中第$(i,\ j)$个元素为 $\sum_{k=1}^{n}a_{ik}b_{kj}$。注意，一般情况下 $AB\neq BA$。还可以得到$(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}$。

另外一种矩阵的乘积叫作克罗内克(Kronecker)积。两个矩阵 $A\in\mathbb{R}^{m\times n}$，$B\in\mathbb{R}^{n\times p}$ 的克罗内克积记作 $A\otimes B$，矩阵维数为 $mp\times nq$，即