

应用数学译丛

卡尔曼滤波 及其实时应用 (第4版)

[美] Charles K. Chui [中] Guanrong Chen 著
戴洪德 周绍磊 戴邵武 李娟 译
秦永元 审
戴邵武 校



NLIC2970893426

Kalman Filtering with
Real-Time Applications
(Fourth Edition)

清华大学出版社

应用数学译丛

卡尔曼滤波 及其实时应用

(第4版)

[美] Charles K. Chui [中] Guanrong Chen 著

戴洪德 周绍磊 戴邵武 李娟 译



NLIC2970893426

Kalman Filtering with
Real-Time Applications
(Fourth Edition)

清华大学出版社
北京

内 容 简 介

本书将理论和应用相结合,深入浅出地介绍了卡尔曼滤波的基本原理和相关的重要主题。从推导、理解卡尔曼滤波必须具备的数学知识入手,首先给出了卡尔曼滤波的直观理解和严格的正交投影证明;在此基础上,针对卡尔曼滤波在实际应用时遇到的不同问题,介绍了系统噪声和量测噪声相关的卡尔曼滤波、有色噪声的处理方法、时不变系统的极限卡尔曼滤波、序贯算法和平方根算法、非线性系统的扩展卡尔曼滤波、高维系统的解耦卡尔曼滤波、不确定系统的区间卡尔曼滤波、随机信号多分辨分析的小波卡尔曼滤波等,并在最后一章简单列举了主体部分没有介绍到的卡尔曼滤波的一些其他重要主题;最后给出了每一章练习题的解答或提示。

本书可以作为通信、导航、自动化、电子、应用数学等专业高年级本科生或研究生的教学用书,也可作为工程技术人员的参考书。本书也适合自学,任何具备基本线性代数、概率论和系统工程知识的读者都能够理解本书。

Translation from English language edition: Kalman Filtering by Charles K. Chui and Guanrong Chen
Copyright © 2009, Springer Berlin Heidelberg Springer. Berlin Heidelberg is a part of Springer Science+Business Media.

All Rights Reserved.

本书中文简体字翻译版由德国施普林格公司授权清华大学出版社在中华人民共和国境内(不包括中国香港、澳门特别行政区和中国台湾地区)独家出版发行。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

北京市版权局著作权合同登记号 图字: 01-2012-3837

版权所有,侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

卡尔曼滤波及其实时应用: 第4版/(美)崔锦泰,(中)陈关荣著; 戴洪德等译. --北京: 清华大学出版社, 2013.4

书名原文: Kalman Filtering with Real-Time Applications(4th ed)

(应用数学译丛)

ISBN 978-7-302-30907-9

I. ①卡… II. ①崔… ②陈… ③戴… III. ①卡尔曼滤波—研究 IV. ①O211.64

中国版本图书馆 CIP 数据核字(2012)第 291843 号

责任编辑: 石磊 赵从棉

封面设计: 傅瑞学

责任校对: 刘玉霞

责任印制: 刘海龙

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社总机: 010-62770175 邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者: 清华大学印刷厂

经 销: 全国新华书店

开 本: 170mm×230mm 印 张: 13.25 字 数: 245 千字

版 次: 2013 年 4 月第 1 版 印 次: 2013 年 4 月第 1 次印刷

印 数: 1~3000

定 价: 32.00 元

产品编号: 044397-01

译者序

PREFACE

卡尔曼滤波自从 20 世纪 60 年代初问世以来,就在航空航天领域获得了非常成功的应用。随着研究的深入,卡尔曼滤波技术越来越多地应用于各个领域,如导航制导、工业控制、目标跟踪、大地测量和金融等。

大部分初学者认为卡尔曼滤波理论性较强,门槛较高,且国内专门针对卡尔曼滤波的专著相对较少,给初学者带来了较大的不便。译者在求学时有幸拜读了 Charles K. Chui 和 Guanrong Chen 两位教授的专著《Kalman Filtering with Real-Time Applications》的第 3 版,获益匪浅。2009 年,我们欣喜地发现时隔 11 年后,两位教授推出了这本专著的第 4 版,我们觉得很有必要把这本从出版到现在经过 25 年时间考验和反复修订的专著介绍给国内的读者。本书深入浅出地介绍了卡尔曼滤波的基本原理和发展,以及它们的实时应用。

本书由海军航空工程学院的戴洪德、周绍磊、戴邵武和鲁东大学的李娟合作翻译,另外海军航空工程学院电子信息工程系的吴芳教员参与了第 11 章的翻译,博士研究生李飞参与了第 7、8 章的翻译,尹高阳参与了第 4 章的翻译。参加翻译的还有徐庆九、于进勇、吴光彬、吴晓男、支岳、曹文静、徐胜红、丛源材、吴青坡、蒋华、袁锐、赵伟等。全书由戴洪德统稿,秦永元主审,戴邵武校对。

感谢西北工业大学自动化学院的陈明教授,把我带进卡尔曼滤波的奇妙世界,并把这本书的原著介绍给我。非常感谢西北工业大学的秦永元教授百忙中对译稿进行了非常认真细致的审校。秦永元教授是研究卡尔曼滤波的著名专家,他在 1998 年出版的《卡尔曼滤波与组合导航原理》是该领域的经典著作,成为相关领域研究者案头必备的参考书,并于 2012 年 6 月进行了再版。感谢西北工业大学的严恭敏副教授在本书翻译过程中与译者的深入交流和有益建议。

非常感谢清华大学出版社石磊和赵从棉两位老师为本书的顺利出版所付出的辛勤劳动。

在出版清样第一稿完成后,非常荣幸地联系上了原著作者之一陈关荣教授。陈教授在百忙之中对译稿进行了校对,特此致谢。

虽然经过了反复的校对和修改,由于译者水平有限,难免存在不足之处,恳请读者批评指正。

戴洪德

2012 年 12 月

于烟台海军航空工程学院

第3版前言

----- FOREWORD -----



第3版的《卡尔曼滤波及其实时应用》中增加了两个关于卡尔曼滤波的新主题。为了扩展卡尔曼滤波在不确定系统中的应用,增加了区间卡尔曼滤波(第10章);结合高效的小波和样条技术,介绍了小波卡尔曼滤波(第11章),并给出了信号估计和信号分解等领域内更为有效的计算方案。希望加入这两章能使新版本给出更完整和与时俱进的实时应用卡尔曼滤波技术。

Charles K. Chui
Guanrong Chen
1998.8

第2版前言

----- FOREWORD -----



除了进行较少的勘误和参考文献更新外,我们将第1版第10章的“实时系统辨识”扩展为两节,合并到第8章。在第10章包含了非常基本的小波分析介绍。虽然小波分解和重构的金字塔算法和卡尔曼滤波算法截然不同,它们仍然可以应用到时域滤波。希望在不久的将来,样条和小波可以与卡尔曼滤波相结合。

Charles K. Chui

Guanrong Chen

1990. 9

第1版前言

FOREWORD



卡尔曼滤波作为一种最优状态估计方法,可以应用于受随机干扰的动态系统。准确地说,卡尔曼滤波器给出了一种递推算法,由实时获得的受噪声污染的离散观测数据,对系统状态进行线性、无偏及最小误差方差的最优估计。该算法已经广泛应用于工业和控制的许多领域,如视频和激光跟踪系统、卫星导航、弹道导弹轨迹估计、雷达和火力控制等。随着最近高速计算机的发展,卡尔曼滤波的应用将更加广泛,特别是在更加复杂的实时应用中。

尽管如此重要,卡尔曼滤波的数学理论以及含义并没有被很好地理解,甚至对于一些应用数学家和工程师也是如此。事实上,非常多的应用者仅仅被告知滤波算法是什么,而不知道为什么它们如此有效。本书的一个主要目标就是通过对卡尔曼滤波的数学理论和多种实时应用问题的讨论,来揭开卡尔曼滤波器的神秘面纱。

本书首先介绍了滤波方程的基本推导。该方案的优势是通过假设某些矩阵非奇异,可以很好地理解卡尔曼滤波的最优性。当然通过应用广为人知的正交投影方法,有时也称之为新息方法,可以不需要这些假设。然后将该方法进行扩展来处理系统和量测噪声相关的问题,以及有色噪声问题。本书还讨论了针对非线性系统的卡尔曼滤波及其在自适应系统辨识中的应用。此外,介绍了实时应用中的极限或稳态卡尔曼滤波理论、序贯算法和平方根算法等高效计算方法。卡尔曼滤波一个典型的应用是数字跟踪滤波器设计,如 $\alpha-\beta-\gamma$ 和 $\alpha-\beta-\gamma-\theta$ 跟踪器。对于白噪声,应用卡尔曼增益的极限值来定义 α 、 β 和 γ 参数,对于有色噪声则为 α 、 β 、 γ 和 θ ,可以将该跟踪滤波器描述为极限或稳态卡尔曼滤波器。因为最优估计的误差随着时间以指数衰减,从这个角度看,通过这些更有效的预测-校正方程得到的状态估计是近似最优的。我们还研究了一种可以得到状态向量各分量滤波方程的解耦方法。

本书的写作风格趋向于随意而非刻板式的,数学证明趋向于简单但是严谨,使得具备基本的线性代数和系统理论的任何人,无论是学生还是专家,都易于阅读。考虑到这一点,本书引入了一个预备知识章节,包含了矩阵理论、行列式、概率论和最小二乘原理。为了说明相关知识点,加强对材料的理解,或完成书中的一些证明,在每一章都配备了一定数量的练习题,并在书的末尾给出了答案或提

示。为了满足感兴趣读者进一步的研究,附录材料和参考文献也列在书尾。

本书的设计是为了达到三个目的:不仅适用于自学;而且适用于应用数学或工程专业大学高年级学生或一年级研究生的半学期或一学期的卡尔曼滤波课程;另外,希望本书能够成为工业或控制工程师有价值的参考资料。

第一作者要感谢美国军队研究办公室的持续资助,特别感谢白沙导弹靶场(White Sands Missile Range)的Robert Green的鼓励和多次深入的讨论。对于爱妻,Margaret,作者要感谢她的理解和一如既往的支持。第二作者非常感谢中山大学的陈铭俊教授将这一非常重要的研究领域介绍给作者,以及感谢他夫人Qiyun Xian的耐心和鼓励。

在给出有价值建议的同事中,作者要特别感谢Andrew Chan教授(得克萨斯A&M)、Thomas Huang教授(伊利诺斯)、Tomas Kailath教授(斯坦福)。最后,非常感谢Helmut Lotsch博士和Angela Lahee博士友好的合作和帮助,以及Springer-Verlag编辑们的工作。

Charles K. Chui
Guanrong Chen
1987.1

符号说明

| | |
|------------------------------|--------------------------------------|
| \mathbf{A}, \mathbf{A}_k | 系统矩阵 |
| \mathbf{A}^c | 矩阵 \mathbf{A} 在 Cholesky 分解时的“平方根” |
| \mathbf{A}^u | 矩阵 \mathbf{A} 在上三角分解时的“平方根” |
| \mathbf{B}, \mathbf{B}_k | 控制输入矩阵 |
| \mathbf{C}, \mathbf{C}_k | 量测矩阵 |
| $\text{Cov}(X, Y)$ | 随机变量 X 和 Y 的协方差 |
| $E(X)$ | 随机变量 X 的期望 |
| $E(X Y=y)$ | 条件期望 |
| e_j, \hat{e}_j | |
| $f(x)$ | 概率密度函数 |
| $f(x_1, x_2)$ | 联合概率密度函数 |
| $f(x_1 x_2)$ | 条件概率密度函数 |
| $f_k(\mathbf{x}_k)$ | 非线性向量函数 |
| \mathbf{G} | 极限卡尔曼增益矩阵 |
| \mathbf{G}_k | 卡尔曼增益矩阵 |
| $\mathbf{H}_k(\mathbf{x}_k)$ | 非线性矩阵函数 |
| \mathbf{H}^* | |
| \mathbf{I}_n | $n \times n$ 单位矩阵 |
| \mathbf{J} | 矩阵的若尔当标准型 |
| \mathbf{K}_k | |
| $L(\mathbf{x}, \mathbf{v})$ | |
| \mathbf{M}_{AR} | 可控性矩阵 |
| \mathbf{N}_{CA} | 可观测性矩阵 |
| $\mathbf{O}_{n \times m}$ | $n \times m$ 阶零矩阵 |
| \mathbf{P} | 极限(误差)协方差矩阵 |
| $\mathbf{P}_{k,k}$ | 估计(误差)协方差矩阵 |
| $\mathbf{P}[i,j]$ | 矩阵 \mathbf{P} 的第 (i,j) 个元素 |

| | |
|--|--|
| $P(X)$ | 随机变量 X 的概率 |
| Q_k | 随机向量 ξ_k 的方差矩阵 |
| R_k | 随机向量 η_k 的方差矩阵 |
| \mathbf{R}^n | 列向量空间 $x = [x_1 \quad \cdots \quad x_n]^T$ |
| S_k | ξ_k 和 η_k 的协方差矩阵 |
| tr | 迹 |
| u_k | 在第 k 时刻的确定性控制输入 |
| $\text{Var}(X)$ | 随机变量 X 的方差 |
| $\text{Var}(X Y=y)$ | 条件方差 |
| v_k | 观测(量测)数据 |
| $v^{2\#}$ | |
| W_k | 权值矩阵 |
| w_j | |
| $(W_\phi f)(b, a)$ | 小波积分变换 |
| x_k | k 时刻的状态向量 |
| $\hat{x}_k, \hat{x}_{k k}$ | x_k 的最优滤波估计 |
| $\hat{x}_{k k-1}$ | x_k 的最优预测 |
| \dot{x}_k | x_k 的次优滤波估计 |
| \vec{x}_k | x_k 的近优滤波估计 |
| x^* | |
| $x^\#, x_k^\#$ | |
| $\ w\ $ | w 的范数 |
| $\langle x, w \rangle$ | x, w 的内积 |
| $Y(w_0, \dots, w_r)$ | 向量 w_0, \dots, w_r 的线性生成空间 |
| $\{z_j\}$ | 数据的新息序列 |
| $\alpha, \beta, \gamma, \theta$ | 跟踪器参数 |
| $\{\underline{\beta}_k\}, \{\underline{\gamma}_k\}$ | 白噪声序列 |
| Γ, Γ_k | 系统噪声矩阵 |
| δ_{ij} | Kronecker 符号 |
| $\underline{\varepsilon}_{k,l}, \bar{\varepsilon}_{k,l}$ | 随机(噪声)向量 |
| $\underline{\eta}_k$ | (k 时刻的)量测噪声 |
| $\underline{\xi}_k$ | (k 时刻的)系统噪声 |
| Φ_k | 转移矩阵 |
| df/dA | 雅克比矩阵 |
| $\partial h/\partial x$ | 雅克比矩阵 |

目 录

CONTENTS

| | |
|-------------------------------|----|
| 第 1 章 预备知识 | 1 |
| 1.1 矩阵和行列式初步 | 1 |
| 1.2 概率论初步 | 6 |
| 1.3 最小二乘初步 | 12 |
| 练习 | 14 |
| 第 2 章 卡尔曼滤波：简单推导 | 16 |
| 2.1 模型 | 16 |
| 2.2 最优准则 | 17 |
| 2.3 预测-校正公式 | 19 |
| 2.4 卡尔曼滤波过程 | 22 |
| 练习 | 23 |
| 第 3 章 正交投影和卡尔曼滤波 | 27 |
| 3.1 最优估计的正交性 | 27 |
| 3.2 新息序列 | 29 |
| 3.3 最小方差估计 | 31 |
| 3.4 卡尔曼滤波方程 | 31 |
| 3.5 实时跟踪 | 36 |
| 练习 | 38 |
| 第 4 章 系统噪声和量测噪声相关的卡尔曼滤波 | 42 |
| 4.1 仿射模型 | 42 |
| 4.2 最优估计算子 | 44 |
| 4.3 额外数据对最优估计的影响 | 45 |
| 4.4 卡尔曼滤波方程推导 | 47 |
| 4.5 实时应用 | 52 |

| | |
|-------------------------------|-----------|
| 4.6 线性确定/随机系统 | 54 |
| 练习 | 55 |
| 第5章 有色噪声 | 58 |
| 5.1 处理思路 | 58 |
| 5.2 误差估计 | 59 |
| 5.3 卡尔曼滤波过程 | 61 |
| 5.4 系统白噪声 | 63 |
| 5.5 实时应用 | 64 |
| 练习 | 66 |
| 第6章 极限(稳态)卡尔曼滤波 | 68 |
| 6.1 处理思路 | 69 |
| 6.2 主要结论 | 70 |
| 6.3 几何收敛 | 77 |
| 6.4 实时应用 | 81 |
| 练习 | 82 |
| 第7章 序贯算法和平方根算法 | 84 |
| 7.1 序贯算法 | 84 |
| 7.2 平方根算法 | 88 |
| 7.3 实时应用算法 | 91 |
| 练习 | 91 |
| 第8章 扩展卡尔曼滤波和系统辨识 | 93 |
| 8.1 扩展卡尔曼滤波 | 93 |
| 8.2 卫星轨道估计 | 96 |
| 8.3 自适应系统辨识 | 97 |
| 8.4 一个常值参数辨识的例子 | 100 |
| 8.5 改进的扩展卡尔曼滤波 | 102 |
| 8.6 时变参数辨识 | 107 |
| 练习 | 111 |

| | |
|---|-----|
| 第 9 章 滤波方程解耦 | 113 |
| 9.1 解耦公式 | 113 |
| 9.2 实时跟踪 | 115 |
| 9.3 α - β - γ 跟踪器 | 117 |
| 9.4 一个例子 | 120 |
| 练习 | 121 |
| 第 10 章 区间系统的卡尔曼滤波 | 124 |
| 10.1 区间数学 | 125 |
| 10.1.1 区间及其特性 | 125 |
| 10.1.2 区间运算 | 126 |
| 10.1.3 有理区间函数 | 129 |
| 10.1.4 区间期望和方差 | 130 |
| 10.2 区间卡尔曼滤波 | 132 |
| 10.2.1 区间卡尔曼滤波方案 | 133 |
| 10.2.2 次优区间卡尔曼滤波 | 134 |
| 10.2.3 目标跟踪的例子 | 135 |
| 10.3 加权平均区间卡尔曼滤波 | 137 |
| 练习 | 138 |
| 第 11 章 小波卡尔曼滤波 | 141 |
| 11.1 小波初步 | 141 |
| 11.1.1 小波基础 | 142 |
| 11.1.2 离散小波变换和滤波器组 | 143 |
| 11.2 信号估计和分解 | 145 |
| 11.2.1 随机信号的估计和分解 | 146 |
| 11.2.2 一个随机游走的例子 | 149 |
| 练习 | 151 |
| 第 12 章 附录 | 152 |
| 12.1 卡尔曼平滑器 | 152 |
| 12.2 α - β - γ - θ 跟踪器 | 154 |
| 12.3 自适应卡尔曼滤波 | 156 |

| | | |
|---------|-------------------|-----|
| 12.4 | 自适应卡尔曼滤波在维纳滤波中的应用 | 157 |
| 12.5 | 卡尔曼-布希滤波 | 158 |
| 12.6 | 随机最优控制 | 159 |
| 12.7 | 平方根滤波及其脉动阵列实现 | 160 |
| 参考文献 | | 163 |
| 练习答案和提示 | | 167 |
| 索引 | | 190 |

第 1 章

预备知识

卡尔曼滤波器在工程应用中的重要性得到了广泛的认可，并且建立了严格的数学理论。本书的主要目的是全面讨论卡尔曼滤波的数学理论、计算算法及其在实时跟踪问题中的应用。

为了解释如何得到卡尔曼滤波算法，以及它有哪些优良的性能，必须应用一些矩阵理论的公式和不等式。此外，在实时应用中考虑了系统和量测噪声的统计特性，必须具备一些概率论的基本概念。本章将专门研究这些主题。

1.1 矩阵和行列式初步

用 \mathbf{R}^n 表示所有列向量 $x = [x_1 \quad \cdots \quad x_n]^T$ 的空间，其中 x_1, \dots, x_n 是实数。对于所有 \mathbf{R}^n 中的非零向量 x ，如果 $x^T A x$ 是一个正数，则称 $n \times n$ 的实矩阵 A 是正定的；如果 $x^T A x$ 非负，则称 A 是非负定的。如果 A 和 B 是两个任意 $n \times n$ 阶实数矩阵，当 $A - B$ 正定时，可以表示为

$$A > B$$

当 $A - B$ 非负定时，可以表示为

$$A \geq B$$

首先来复习施瓦兹不等式 (Schwarz inequality)：

$$|x^T y| \leq \|x\| \|y\|, \quad x, y \in \mathbf{R}^n$$

式中， $|x| = (x^T x)^{1/2}$ 。另外，当且仅当 x, y 平行时，上面的不等式变成等式，即

$$x = \lambda y \quad \text{或} \quad y = \lambda x$$

λ 为比例因子。特别地,如果 $y \neq 0$,施瓦兹不等式可以写为

$$x^T x \geq (y^T x)^T (y^T y)^{-1} (y^T x)$$

借助这个公式,可以将施瓦兹不等式推广到矩阵形式。

引理 1.1(矩阵施瓦兹不等式(Matrix Schwarz inequality)) 若 P 和 Q 分别是 $m \times n$ 和 $m \times l$ 矩阵, $P^T P$ 非奇异, 则

$$Q^T Q \geq (P^T Q)^T (P^T P)^{-1} (P^T Q) \quad (1.1)$$

此外,对于某些 $m \times l$ 矩阵 S , 当且仅当 $Q = PS$ 时, 式(1.1)的等号成立。 ■

证明如下。

向量形式的施瓦兹不等式的证明比较简单。二次多项式^①

$$(x - \lambda y)^T (x - \lambda y) \geq 0, \quad y \neq 0$$

关于 λ 的最小值在

$$\lambda = (y^T y)^{-1} (y^T x)$$

时得到。将该 λ 值代入上面的不等式, 即得施瓦兹不等式。在矩阵形式中, 考虑到

$$(Q - PS)^T (Q - PS) \geq 0$$

同时令

$$S = (P^T P)^{-1} (P^T Q)$$

得到

$$Q^T Q \geq S^T (P^T Q) + (P^T Q)^T S - S^T (P^T P) S = (P^T Q)^T (P^T P)^{-1} (P^T Q)$$

和式(1.1)相同。对于 $n \times l$ 矩阵 S , 当且仅当

$$(Q - PS)^T (Q - PS) = 0$$

即 $Q = PS$ 时, 该不等式的等号成立。这就完成了该引理的证明。

下面来看矩阵求逆引理:

引理 1.2(矩阵求逆引理(Matrix inversion lemma)) 令

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

式中, \mathbf{A}_{11} 和 \mathbf{A}_{22} 分别是 $n \times n$ 和 $m \times m$ 的非奇异子矩阵。这样,

$$(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}) \text{ 和 } (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})$$

都是非奇异的, 所以 \mathbf{A} 非奇异,

$$\begin{aligned} \mathbf{A}^{-1} &= \begin{bmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \\ -(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & -(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{bmatrix}. \end{aligned} \quad (1.2)$$

^① 下式译者作了微小修订, 补充了“ ≥ 0 ”。

特别地,

$$(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} = \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} \quad (1.3)$$

及

$$\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} = (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \quad (1.4)$$

更进一步有

$$\begin{aligned} \det \mathbf{A} &= (\det \mathbf{A}_{11}) \det (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}) \\ &= (\det \mathbf{A}_{22}) \det (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}) \end{aligned} \quad (1.5) \blacksquare$$

证明如下。

可以将 \mathbf{A} 表示为

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix}$$

或

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_n & \mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

取行列式, 得到(1.5)式。特别地, 从这里我们获得

$$\det \mathbf{A} \neq 0$$

即 \mathbf{A} 是非奇异的。接下来, 注意到

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \\ \mathbf{0} & (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \end{bmatrix}$$

及

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_m \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_m \end{bmatrix}$$

可得

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_m \end{bmatrix}^{-1}$$

这样就可以得到式(1.2)的第一部分。式(1.2)的第二部分可以用同样的方法证明。式(1.3)和式(1.4)可以根据式(1.2)的相应矩阵块相等得到。

直接应用引理 1.2 可以得到下面的结果。

引理 1.3 如果 $\mathbf{P} \geqslant \mathbf{Q} > \mathbf{0}$, 则 $\mathbf{Q}^{-1} \geqslant \mathbf{P}^{-1} > \mathbf{0}$ 。 ■

证明如下。

令 $\mathbf{P}(\epsilon) = \mathbf{P} + \epsilon \mathbf{I}$, 其中 $\epsilon > 0$, 则 $\mathbf{P}(\epsilon) - \mathbf{Q} > \mathbf{0}$ 。根据引理 1.2, 有

$$\begin{aligned}\mathbf{P}^{-1}(\epsilon) &= [\mathbf{Q} + (\mathbf{P}(\epsilon) - \mathbf{Q})]^{-1} \\ &= \mathbf{Q}^{-1} - \mathbf{Q}^{-1}[(\mathbf{P}(\epsilon) - \mathbf{Q})^{-1} + \mathbf{Q}^{-1}]^{-1} \mathbf{Q}^{-1}\end{aligned}$$

可得

$$\mathbf{Q}^{-1} - \mathbf{P}^{-1}(\epsilon) = \mathbf{Q}^{-1}[(\mathbf{P}(\epsilon) - \mathbf{Q})^{-1} + \mathbf{Q}^{-1}]^{-1} \mathbf{Q}^{-1} > \mathbf{0}$$

令 $\epsilon \rightarrow 0$, 得到 $\mathbf{Q}^{-1} - \mathbf{P}^{-1} \geq \mathbf{0}$, 所以

$$\mathbf{Q}^{-1} \geq \mathbf{P}^{-1} > \mathbf{0}$$

下面讨论 $n \times n$ 矩阵 \mathbf{A} 的迹。矩阵 \mathbf{A} 的迹可以表示为 $\text{tr}\mathbf{A}$, 定义为矩阵 \mathbf{A} 对角线元素的和, 即

$$\text{tr}\mathbf{A} = \sum_{i=1}^n a_{ii}$$

式中, $\mathbf{A} = [a_{ij}]$ 。首先来介绍一些基本性质。

引理 1.4 如果 \mathbf{A} 和 \mathbf{B} 都是 $n \times n$ 阶矩阵, 则

$$\text{tr}\mathbf{A}^T = \text{tr}\mathbf{A} \quad (1.6)$$

$$\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}\mathbf{A} + \text{tr}\mathbf{B} \quad (1.7)$$

$$\text{tr}(\lambda\mathbf{A}) = \lambda\text{tr}\mathbf{A} \quad (1.8)$$

如果 \mathbf{A} 是 $n \times m$ 阶矩阵, \mathbf{B} 是 $m \times n$ 阶矩阵, 则

$$\text{tr}\mathbf{AB} = \text{tr}\mathbf{B}^T\mathbf{A}^T = \text{tr}\mathbf{BA} = \text{tr}\mathbf{A}^T\mathbf{B}^T \quad (1.9)$$

$$\text{tr}\mathbf{A}^T\mathbf{A} = \sum_{i=1}^n \sum_{j=1}^m a_{ij}^2 \quad (1.10) \blacksquare$$

上面恒等式的证明可以从定义直接得到, 作为练习(见练习 1.1)留给读者自己完成。下面的结论很重要。

引理 1.5 令 $n \times n$ 阶矩阵 \mathbf{A} 的特征值是 $\lambda_1, \dots, \lambda_n$, 重根也列入其中, 则

$$\text{tr}\mathbf{A} = \sum_{i=1}^n \lambda_i \quad (1.11) \blacksquare$$

证明如下。

将 \mathbf{A} 写为 $\mathbf{A} = \mathbf{U}\mathbf{J}\mathbf{U}^{-1}$, 其中 \mathbf{J} 是矩阵 \mathbf{A} 的若尔当标准型(Jordan canonical form), \mathbf{U} 为非奇异矩阵。应用式(1.9)可以得到

$$\text{tr}\mathbf{A} = \text{tr}(\mathbf{AU})\mathbf{U}^{-1} = \text{tr}\mathbf{U}^{-1}(\mathbf{AU}) = \text{tr}\mathbf{J} = \sum_{i=1}^n \lambda_i$$

根据这个引理, 如果 $\mathbf{A} > \mathbf{0}$, 则 $\text{tr}\mathbf{A} > 0$; 如果 $\mathbf{A} \geq \mathbf{0}$, 则 $\text{tr}\mathbf{A} \geq 0$ 。

下面介绍一些关于迹的有用不等式。

引理 1.6 令 \mathbf{A} 为 $n \times n$ 阶矩阵, 则

$$\text{tr}\mathbf{A} \leq (n\text{tr}\mathbf{AA}^T)^{1/2} \quad (1.12) \blacksquare$$

上面不等式的证明作为练习(见练习 1.2)留给读者自己完成。

引理 1.7 如果 A 和 B 分别是 $n \times m$ 和 $m \times l$ 阶矩阵, 则

$$\operatorname{tr}(AB)(AB)^T \leqslant (\operatorname{tr}AA^T)(\operatorname{tr}BB^T)$$

相应的对于任意适当维数的矩阵 A_1, \dots, A_p ,

$$\operatorname{tr}(A_1 \cdots A_p)(A_1 \cdots A_p)^T \leqslant (\operatorname{tr}A_1 A_1^T) \cdots (\operatorname{tr}A_p A_p^T) \quad (1.13) \blacksquare$$

证明如下。

如果 $A = [a_{ij}]$, $B = [b_{ij}]$, 根据施瓦兹不等式, 则

$$\begin{aligned} \operatorname{tr}(AB)(AB)^T &= \operatorname{tr} \left[\sum_{k=1}^m a_{ik} b_{kj} \right] \left[\sum_{k=1}^m a_{jk} b_{ki} \right] \\ &= \operatorname{tr} \left[\sum_{p=1}^l \left(\sum_{k=1}^m a_{ik} b_{kp} \right)^2 \right. \\ &\quad \left. \ddots \right. \\ &\quad \left. * \quad \sum_{p=1}^l \left(\sum_{k=1}^m a_{nk} b_{kp} \right)^2 \right] \\ &= \sum_{i=1}^n \sum_{p=1}^l \left(\sum_{k=1}^m a_{ik} b_{kp} \right)^2 \leqslant \sum_{i=1}^n \sum_{p=1}^l \sum_{k=1}^m a_{ik}^2 \sum_{k=1}^m b_{kp}^2 \\ &= \left(\sum_{i=1}^n \sum_{k=1}^m a_{ik}^2 \right) \left(\sum_{p=1}^l \sum_{k=1}^m b_{kp}^2 \right) \\ &= (\operatorname{tr}AA^T)(\operatorname{tr}BB^T) \end{aligned}$$

这就完成了该引理的证明。

必须注意, 对于 $A \geqslant B > 0$, 不一定就有 $\operatorname{tr}AA^T \geqslant \operatorname{tr}BB^T$, 一个例子为

$$A = \begin{bmatrix} 12 & 0 \\ 5 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$$

很显然, $A - B \geqslant 0$ 和 $B > 0$, 但是

$$\operatorname{tr}AA^T = \frac{169}{25} < 7 = \operatorname{tr}BB^T$$

(见练习 1.3)。

对于对称矩阵, 可以用下面的方法得到需要的结论。

引理 1.8 令 A 和 B 都是非负定对称矩阵, 且 $A \geqslant B$, 则 $\operatorname{tr}AA^T \geqslant \operatorname{tr}BB^T$, 或 $\operatorname{tr}A^2 \geqslant \operatorname{tr}B^2$ 。 \blacksquare

将该引理的证明留作练习(见练习 1.4)。

引理 1.9 令 B 是 $n \times n$ 阶非负定对称矩阵, 则

$$\operatorname{tr}B^2 \leqslant (\operatorname{tr}B)^2 \quad (1.14)$$

若 A 是另一个 $n \times n$ 阶非负定对称矩阵, 且 $B \leqslant A$, 则

$$\text{tr} \mathbf{B}^2 \leq (\text{tr} \mathbf{A})^2 \quad (1.15) \blacksquare$$

为了证明式(1.14),令 \mathbf{B} 的特征值为 $\lambda_1, \dots, \lambda_n$,则 \mathbf{B}^2 的特征值为 $\lambda_1^2, \dots, \lambda_n^2$ 。因为 $\lambda_1, \dots, \lambda_n$ 是非负的,根据引理 1.5 有

$$\text{tr} \mathbf{B}^2 = \sum_{i=1}^n \lambda_i^2 \leq \left(\sum_{i=1}^n \lambda_i \right)^2 = (\text{tr} \mathbf{B})^2$$

又 $\mathbf{B} \leq \mathbf{A}$,则 $\text{tr} \mathbf{B} \leq \text{tr} \mathbf{A}$,可得式(1.15)。

还有下面在以后章节中非常有用结论。

引理 1.10 设 \mathbf{F} 是一个特征值为 $\lambda_1, \dots, \lambda_n$ 的 $n \times n$ 阶矩阵。若

$$\lambda := \max(|\lambda_1|, \dots, |\lambda_n|) < 1$$

则存在满足 $0 < r < 1$ 的实数 r 和常数 C ,使得

$$|\text{tr} \mathbf{F}^k (\mathbf{F}^k)^T| \leq C r^k, \quad k = 1, 2, \dots \quad \blacksquare$$

证明如下。

令 \mathbf{J} 是矩阵 \mathbf{F} 的若尔当标准型,则存在非奇异矩阵 \mathbf{U} ,使 $\mathbf{F} = \mathbf{U} \mathbf{J} \mathbf{U}^{-1}$ 。根据式(1.13),得

$$\begin{aligned} |\text{tr} \mathbf{F}^k (\mathbf{F}^k)^T| &= |\text{tr} \mathbf{U} \mathbf{J}^k \mathbf{U}^{-1} (\mathbf{U}^{-1})^T (\mathbf{J}^k)^T \mathbf{U}^T| \\ &\leq |\text{tr} \mathbf{U} \mathbf{U}^T| |\text{tr} \mathbf{J}^k (\mathbf{J}^k)^T| |\text{tr} \mathbf{U}^{-1} (\mathbf{U}^{-1})^T| \\ &\leq p(k) \lambda^{2k} \end{aligned}$$

式中, $p(k)$ 是 k 的多项式。

取任意的 r 满足 $\lambda^2 < r < 1$,并对所有的 k ,选择正常数 C 满足

$$p(k) \left(\frac{\lambda^2}{r} \right)^k \leq C$$

就可以得到所需结果。

1.2 概率论初步

投掷均匀硬币的试验,在每次投掷时,结果为正面朝上(H),或反面朝上(T)。将试验实际的结果称为试验的输出,所有可能的输出集合称为试验的样本空间(S)。例如,如果投掷一枚均匀的硬币两次,则两次投掷的结果可能为 HH 、 TT 、 HT 或 TH ,集合 $\{HH, TT, HT, TH\}$ 是一个样本空间 S 。此外,该样本的任意子集称为一个事件;如果某个事件只有一个可能的结果,则称为基本事件。

因为没有预测输出结果的方法,需要针对每一个事件定义一个 0 和 1 之间的实数 P ,来描述某个确定事件发生的概率。这可以由一个实值函数来描述,称

为定义在样本空间的随机变量。在上面的例子中,如果随机变量 $X=X(s), s \in S$, 表示 s 次试验中正面朝上的次数(H 发生),则 $P=P(X(s))$ 给出了 s 次试验中正面朝上的概率。令 S 为样本空间, $X: S \rightarrow \mathbf{R}^1$ 为随机变量。对于任何一个可测集合 $A \subset \mathbf{R}^1$ (在本例中, $A=\{0\}, \{1\}, \{2\}$), 分别表示没有出现正面、出现一次正面、出现两次正面), 定义一个 $P: \{\text{事件}\} \rightarrow [0, 1]$, 其中每个事件是一个集合 $\{s \in S: X(s) \in A \subset \mathbf{R}^1\} := \{X \in A\}$, 满足以下条件:

- (1) 对于任意可测集 $A \subset \mathbf{R}^1$, $P(X \in A) \geq 0$;
- (2) $P(X \in \mathbf{R}^1) = 1$;
- (3) 对 \mathbf{R}^1 中的任意两两互不相容的可测集 A_i

$$P\left(X \in \bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(X \in A_i)$$

则称 P 为随机变量 X 的概率分布(或者概率分布函数)。

如果存在一个非负可积函数^① f , 使得

$$P(X \in A) = \int_A f(x) dx \quad (1.16)$$

对所有的可测集 A 成立, 则称 P 为连续概率分布, f 为随机变量 X 的概率密度函数。注意到我们已经定义 $f(x)dx = d\lambda$, 其中 λ 是一个测度(比如阶梯函数), 那么投掷硬币的离散型情形也包含在其中了。

如果相应的概率密度函数 f 定义为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad \sigma > 0, \mu \in \mathbf{R} \quad (1.17)$$

则称为高斯(或者正态)概率密度函数, P 为随机变量 X 的正态分布, 记为 $X \sim N(\mu, \sigma^2)$, 可以简单地证明正态分布 P 为一个概率分布。实际上,(1) 因为 $f(x) > 0$, 对任意可测集 $A \subset \mathbf{R}$, 有 $P(X \in A) = \int_A f(x) dx \geq 0$; (2) 通过变

换 $y = (x - \mu)/(\sqrt{2}\sigma)$, 得

$$P(X \in \mathbf{R}^1) = \int_{-\infty}^{\infty} f(x) dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-y^2} dy = 1$$

(见练习题 1.5); (3) 因为

$$\int_{\bigcup_i A_i} f(x) dx = \sum_i \int_{A_i} f(x) dx$$

对任意两两互不相交的可测集 $A_i \subset \mathbf{R}^1$ 有

^① 译者注: 原文中为“可积函数”。

$$P\left(X \in \bigcup_i A_i\right) = \sum_i P(X \in A_i)$$

令 X 为随机变量, 则 X 的数学期望, 即 X 值的均值, 可以定义为

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad (1.18)$$

对任意以 f 为概率密度函数的随机变量 X , $E(X)$ 为实数。对于正态分布, 再次做变换 $y = (x - \mu)/(\sqrt{2}\sigma)$, 可得

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} xf(x)dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} xe^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\sqrt{2}\sigma y + \mu) e^{-y^2} dy \\ &= \mu \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-y^2} dy \\ &= \mu \end{aligned} \quad (1.19)$$

$E(X)$ 也称为概率密度函数 f 的一阶矩。二阶矩给出了 X 的方差, 定义为

$$\text{Var}(X) = E(X - E(X))^2 = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)dx \quad (1.20)$$

这个值表示随机变量 X 偏离其数学期望 $E(X)$ 的离散程度。对于正态分布, 再作变换 $y = (x - \mu)/(\sqrt{2}\sigma)$, 代入等式 $\int_{-\infty}^{\infty} y^2 e^{-y^2} dy = \sqrt{\pi}/2$ (见练习 1.6), 可得

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2} dy \\ &= \sigma^2 \end{aligned} \quad (1.21)$$

下面来看由随机变量组成的随机向量。记 n 维随机向量 $\mathbf{X} = [X_1 \quad \cdots \quad X_n]^T$, 其中 $X_i(s) \in \mathbf{R}^1, s \in S$ 。

令 P 表示 \mathbf{X} 的连续概率分布函数, 则

$$P(X_1 \in A_1, \dots, X_n \in A_n) = \int_{A_1} \cdots \int_{A_n} f(x_1, \dots, x_n) dx_1 \cdots dx_n \quad (1.22)$$

其中 A_1, \dots, A_n 为 \mathbf{R}^1 上的可测集, f 为可积函数, 则 f 称为随机向量 \mathbf{X} 的联合概率密度函数, P 为随机向量 \mathbf{X} 的联合概率分布(函数)。对于每个 $i(i=1, \dots, n)$, 定义

$$f_i(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n \quad (1.23)$$

显然 $\int_{-\infty}^{\infty} f_i(x) dx = 1$, f_i 为随机向量 \mathbf{X} 的联合概率密度函数 $f(x_1, \dots, x_n)$ 的第 i 个边缘概率密度函数。同理, 按照定义 f_i 的方法, 定义 f_{ij}, f_{ijk} 为分别去掉 x_i, x_j 与 x_i, x_j, x_k 的概率密度函数。如果

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} (\det \mathbf{R})^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \underline{\mu})^T \mathbf{R}^{-1} (\mathbf{x} - \underline{\mu}) \right\} \quad (1.24)$$

式中 $\underline{\mu}$ 是一个 n 维常值向量, \mathbf{R} 为对称正定矩阵, 则 $f(\mathbf{x})$ 是随机向量 \mathbf{X} 的高斯(或正态)概率密度函数。可以证明:

$$\int_{-\infty}^{+\infty} f(\mathbf{x}) d\mathbf{x} := \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(\mathbf{x}) dx_1 \cdots dx_n = 1 \quad (1.25)$$

$$\begin{aligned} E(\mathbf{X}) &= \int_{-\infty}^{+\infty} \mathbf{x} f(\mathbf{x}) d\mathbf{x} \\ &:= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} f(\mathbf{x}) dx_1 \cdots dx_n \\ &= \underline{\mu} \end{aligned} \quad (1.26)$$

及

$$\text{Var}(\mathbf{X}) = E(\mathbf{X} - \underline{\mu})(\mathbf{X} - \underline{\mu})^T = \mathbf{R} \quad (1.27)$$

事实上, 因为 \mathbf{R} 是一个对称正定矩阵, 存在单位矩阵 \mathbf{U} 满足 $\mathbf{R} = \mathbf{U}^T \mathbf{J} \mathbf{U}$, 其中 $\mathbf{J} = \text{diag}[\lambda_1, \dots, \lambda_n]$, 且 $\lambda_1, \dots, \lambda_n > 0$ 。令 $\mathbf{y} = \frac{1}{\sqrt{2}} \text{diag}[\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}] \mathbf{U}(\mathbf{X} - \underline{\mu})$, 则

$$\int_{-\infty}^{+\infty} f(\mathbf{x}) d\mathbf{x} = \frac{2^{n/2} \sqrt{\lambda_1} \cdots \sqrt{\lambda_n}}{(2\pi)^{n/2} (\lambda_1 \cdots \lambda_n)^{1/2}} \int_{-\infty}^{+\infty} e^{-y_1^2} dy_1 \cdots \int_{-\infty}^{+\infty} e^{-y_n^2} dy_n = 1$$

方程(1.26)和方程(1.27)的证明可以应用前面标量情况下采用的代换来完成(见式(1.21)和练习 1.7)。

下面介绍条件概率的概念。考虑这样一个试验, 一个罐子中装有 M_1 个白球和 M_2 个黑球。当采用不放回抽取时, 即第一个球取出后不再放入原来的罐子, 求在第一次取到黑球(事件 A_1)的条件下, 第二次又取到黑球(事件 A_2)的概率。

对于这个简单问题,我们这样解释:因为从罐中取到的第一个球是黑色的,则第二次取球前,罐中剩余 M_1 个白球和 $M_2 - 1$ 个黑球,此时取得黑球的概率为

$$\frac{M_2 - 1}{M_1 + M_2 - 1}$$

又

$$\frac{M_2 - 1}{M_1 + M_2 - 1} = \frac{M_2}{M_1 + M_2} \cdot \frac{M_2 - 1}{M_1 + M_2 - 1} / \frac{M_2}{M_1 + M_2}$$

式中 $\frac{M_2}{M_1 + M_2}$ 为第一次取球时取到黑球的概率, $\frac{M_2}{M_1 + M_2} \cdot \frac{M_2 - 1}{M_1 + M_2 - 1}$ 为第一次和第二次都取到黑球的概率。

根据上述例子,引出条件概率的定义:在 $X_2 \in A_2$ 的条件下, $X_1 \in A_1$ 的条件概率为

$$P(X_1 \in A_1 | X_2 \in A_2) = \frac{P(X_1 \in A_1, X_2 \in A_2)}{P(X_2 \in A_2)} \quad (1.28)$$

假定 P 为联合概率密度函数 f 所对应的连续概率分布函数,则式(1.28)变为

$$P(X_1 \in A_1 | X_2 \in A_2) = \frac{\int_{A_1} \int_{A_2} f(x_1, x_2) dx_1 dx_2}{\int_{A_2} f_2(x_2) dx_2}$$

式中 f_2 定义为

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1$$

为 f 的第二个边缘概率密度函数。

令 $f(x_1 | x_2)$ 表示 $P(X_1 \in A_1 | X_2 \in A_2)$ 所对应的概率密度函数。 $f(x_1 | x_2)$ 被称为条件概率分布函数 $P(X_1 \in A_1 | X_2 \in A_2)$ 所对应的条件概率密度函数。显然,

$$f(x_1 | x_2) = \frac{f(x_1, x_2)}{f_2(x_2)} \quad (1.29)$$

这称作贝叶斯公式(参见《概率论》A. N. Shirayev(1984))。同理,贝叶斯公式也可以写为

$$f(x_1, x_2) = f(x_1 | x_2) f_2(x_2) = f(x_2 | x_1) f_1(x_1) \quad (1.30)$$

上述贝叶斯公式对于随机向量 $\mathbf{X}_1, \mathbf{X}_2$ 也成立。

如果 \mathbf{X} 和 \mathbf{Y} 分别是 n 维和 m 维随机向量,则 \mathbf{X} 和 \mathbf{Y} 的协方差是一个 $n \times m$ 矩阵

$$\text{Cov}(\mathbf{X}, \mathbf{Y}) = E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{Y} - E(\mathbf{Y}))^\top] \quad (1.31)$$

当 $\mathbf{Y} = \mathbf{X}$ 时, 得到方差矩阵, 有时也被称为 \mathbf{X} 的协方差矩阵, $\text{Var}(\mathbf{X}) = \text{Cov}(\mathbf{X}, \mathbf{X})$ 。

可以证明期望、方差、协方差具有以下的性质:

$$\begin{cases} E(\mathbf{AX} + \mathbf{BY}) = \mathbf{AE}(\mathbf{X}) + \mathbf{BE}(\mathbf{Y}) \\ E((\mathbf{AX})(\mathbf{BY})^T) = \mathbf{A}(E(\mathbf{XY}^T))\mathbf{B}^T \\ \text{Var}(\mathbf{X}) \geq \mathbf{0} \\ \text{Cov}(\mathbf{X}, \mathbf{Y}) = (\text{Cov}(\mathbf{Y}, \mathbf{X}))^T \\ \text{Cov}(\mathbf{X}, \mathbf{Y}) = E(\mathbf{XY}^T) - E(\mathbf{X})(E(\mathbf{Y}))^T \end{cases} \quad (1.32)$$

式中 \mathbf{A}, \mathbf{B} 是常值矩阵(见练习 1.8)。

如果 $f(\mathbf{x}|\mathbf{y}) = f_1(\mathbf{x}), f(\mathbf{y}|\mathbf{x}) = f_2(\mathbf{y})$, 则 \mathbf{X} 和 \mathbf{Y} 是独立的; 如果 $\text{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}$, 则称 \mathbf{X} 和 \mathbf{Y} 是不相关的。显然, 如果 \mathbf{X} 和 \mathbf{Y} 是相互独立的, 则 \mathbf{X} 和 \mathbf{Y} 是不相关的。事实上, 如果 \mathbf{X} 和 \mathbf{Y} 是独立的, 则 $f(\mathbf{x}, \mathbf{y}) = f_1(\mathbf{x})f_2(\mathbf{y})$ 。因此

$$\begin{aligned} E(\mathbf{XY}^T) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{xy}^T f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \int_{-\infty}^{\infty} \mathbf{x} f_1(\mathbf{x}) d\mathbf{x} \int_{-\infty}^{\infty} \mathbf{y}^T f_2(\mathbf{y}) d\mathbf{y} \\ &= E(\mathbf{X})(E(\mathbf{Y}))^T \end{aligned}$$

由式(1.32), $\text{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}$ 。但除了在正态概率分布的情形下, 不相关一般不能推出相互独立。令

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim N \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \mathbf{R} \right)$$

式中

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix}, \quad \mathbf{R}_{12} = \mathbf{R}_{21}^T$$

$\mathbf{R}_{11}, \mathbf{R}_{22}$ 是对称的, \mathbf{R} 是正定的。可以证明, 当且仅当 $\mathbf{R}_{12} = \text{Cov}(\mathbf{X}_1, \mathbf{X}_2) = \mathbf{0}$ 时, $\mathbf{X}_1, \mathbf{X}_2$ 是独立的(见练习 1.9)。

如果 \mathbf{X} 和 \mathbf{Y} 是两个随机向量, 与定义期望和方差一样, $\mathbf{Y} = \mathbf{y}$ 条件下 \mathbf{X} 的条件数学期望定义为

$$E(\mathbf{X} | \mathbf{Y} = \mathbf{y}) = \int_{-\infty}^{\infty} \mathbf{x} f(\mathbf{x} | \mathbf{y}) d\mathbf{x} \quad (1.33)$$

而 $\mathbf{Y} = \mathbf{y}$ 条件下, \mathbf{X} 的条件方差定义为

$$\text{Var}(\mathbf{X} | \mathbf{Y} = \mathbf{y}) = \int_{-\infty}^{\infty} [\mathbf{x} - E(\mathbf{X} | \mathbf{Y} = \mathbf{y})][\mathbf{x} - E(\mathbf{X} | \mathbf{Y} = \mathbf{y})]^T f(\mathbf{x} | \mathbf{y}) d\mathbf{x} \quad (1.34)$$

假设

$$E\left(\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}\right) = \begin{bmatrix} \underline{\mu}_x \\ \underline{\mu}_y \end{bmatrix}, \quad \text{Var}\left(\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_{yy} \end{bmatrix}$$

则由式(1.24)可得

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= f\left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}\right) \\ &= \frac{1}{(2\pi)^{n/2} \left(\det \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_{yy} \end{bmatrix}\right)^{1/2}} \\ &\quad \cdot \exp\left\{-\frac{1}{2} \left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} \underline{\mu}_x \\ \underline{\mu}_y \end{bmatrix}\right)^T \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_{yy} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} \underline{\mu}_x \\ \underline{\mu}_y \end{bmatrix}\right)\right\} \end{aligned}$$

可以证明

$$\begin{aligned} f(\mathbf{x} | \mathbf{y}) &= \frac{f(\mathbf{x}, \mathbf{y})}{f(\mathbf{y})} \\ &= \frac{1}{(2\pi)^{n/2} (\det \tilde{\mathbf{R}})^{1/2}} \exp\left\{-\frac{1}{2} (\mathbf{x} - \tilde{\underline{\mu}})^T \tilde{\mathbf{R}}^{-1} (\mathbf{x} - \tilde{\underline{\mu}})\right\} \quad (1.35) \end{aligned}$$

式中

$$\tilde{\underline{\mu}} = \underline{\mu}_x + \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} (\mathbf{y} - \underline{\mu}_y), \quad \tilde{\mathbf{R}} = \mathbf{R}_{xx} - \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx}$$

(见练习 1.10)代入 $\tilde{\underline{\mu}}$ 和 $\tilde{\mathbf{R}}$, 可得

$$E(\mathbf{X} | \mathbf{Y} = \mathbf{y}) = E(\mathbf{X}) + \text{Cov}(\mathbf{X}, \mathbf{Y}) [\text{Var}(\mathbf{Y})]^{-1} (\mathbf{y} - E(\mathbf{Y})) \quad (1.36)$$

$$\text{Var}(\mathbf{X} | \mathbf{Y} = \mathbf{y}) = \text{Var}(\mathbf{X}) - \text{Cov}(\mathbf{X}, \mathbf{Y}) [\text{Var}(\mathbf{Y})]^{-1} \text{Cov}(\mathbf{Y}, \mathbf{X}) \quad (1.37)$$

1.3 最小二乘初步

设 $\{\xi_k\}$ 是一系列的随机向量, 称为随机序列, 其中 $E(\xi_k) = \underline{\mu}_k$, $\text{Cov}(\xi_k, \xi_j) = \mathbf{R}_{kj}$, 则 $\text{Var}(\xi_k) = \mathbf{R}_{kk} := \mathbf{R}_k$ 。如果 $\text{Cov}(\xi_k, \xi_j) = \mathbf{R}_{kj} = \mathbf{R}_k \delta_{kj}$ (式中 $k = j$ 时 $\delta_{kj} = 1$; $k \neq j$ 时 $\delta_{kj} = 0$), 则这个随机序列称为白噪声序列。如果白噪声序列 $\{\xi_k\}$ 的每个 ξ_k 都是正态的, 则 $\{\xi_k\}$ 称为高斯(或标准)白噪声序列。

一个观测数据被噪声污染的线性系统的观测方程为

$$\mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k + \xi_k$$

式中 $\{\mathbf{x}_k\}$ 是状态序列, $\{\mathbf{u}_k\}$ 是控制序列, $\{\mathbf{v}_k\}$ 是数据序列。假设对于每一个 k ,

$q \times n$ 阶常值矩阵 \mathbf{C}_k , $q \times p$ 阶常值矩阵 \mathbf{D}_k 和 p 阶确定性控制向量 \mathbf{u}_k 已经给出。通常 $\{\underline{\xi}_k\}$ 未知, 但假设为零均值高斯白噪声, 即对于 $k, j=1, 2, \dots$, 有 $E(\underline{\xi}_k) = \mathbf{0}$, 且 $E(\underline{\xi}_k \underline{\xi}_j^T) = \mathbf{R}_k \delta_{kj}$, \mathbf{R}_k 是对称正定矩阵。

我们的目的是从信息 $\{v_k\}$ 中获得状态向量 x_k 的最优估计 \hat{x}_k 。如果没有噪声, $\mathbf{z}_k - \mathbf{C}_k \hat{x}_k = \mathbf{0}$, 式中 $\mathbf{z}_k := v_k - \mathbf{D}_k \mathbf{u}_k$, 该线性系统任何时候都有解, 否则对于所有的 y_k , 必须最小化量测误差 $\mathbf{z}_k - \mathbf{C}_k y_k$ 。通常数据都会被噪声污染, 我们要最小化下式:

$$F(y_k, \mathbf{W}_k) = E[(\mathbf{z}_k - \mathbf{C}_k y_k)^T \mathbf{W}_k (\mathbf{z}_k - \mathbf{C}_k y_k)]$$

对于所有 n 维向量 y_k , \mathbf{W}_k 是正定对称的 $q \times q$ 阶矩阵, 称为权值矩阵。也就是, 希望找到 $\hat{y}_k = \hat{y}_k(\mathbf{W}_k)$, 使得

$$F(\hat{y}_k, \mathbf{W}_k) = \min_{y_k} F(y_k, \mathbf{W}_k)$$

另外, 我们希望定义最优权值矩阵 $\hat{\mathbf{W}}_k$ 。为了找到 $\hat{y}_k = \hat{y}_k(\mathbf{W}_k)$, 假设 $(\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)$ 非奇异, 重新整理为

$$\begin{aligned} F(y_k, \mathbf{W}_k) &= E[(\mathbf{z}_k - \mathbf{C}_k y_k)^T \mathbf{W}_k (\mathbf{z}_k - \mathbf{C}_k y_k)] \\ &= E\{[(\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k) \mathbf{y}_k - \mathbf{C}_k^T \mathbf{W}_k \mathbf{z}_k]^T (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} [(\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k) \mathbf{y}_k - \mathbf{C}_k^T \mathbf{W}_k \mathbf{z}_k]\} \\ &\quad + E(\mathbf{z}_k^T [\mathbf{I} - \mathbf{W}_k \mathbf{C}_k (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T] \mathbf{W}_k \mathbf{z}_k) \end{aligned}$$

式中右边的第一项是非负定的。为了最小化 $F(y_k, \mathbf{W}_k)$, 右边的第一项必须为零, 则

$$\hat{y}_k = (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{W}_k \mathbf{z}_k$$

若 $(\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)$ 奇异, 则 \hat{y}_k 不唯一。为了找到最优权值 $\hat{\mathbf{W}}_k$, 考虑

$$F(\hat{y}_k, \mathbf{W}_k) = E[(\mathbf{z}_k - \mathbf{C}_k \hat{y}_k)^T \mathbf{W}_k (\mathbf{z}_k - \mathbf{C}_k \hat{y}_k)]$$

很显然, 对于正定的权值矩阵 \mathbf{W}_k , 该式不存在最小值, 因为最小值在 $\mathbf{W}_k = \mathbf{0}$ 时获得。所以, 需要另外的量测来确定一个最优 $\hat{\mathbf{W}}_k$ 。注意原始问题是根据 $\hat{y}_k(\mathbf{W}_k)$ 估计状态向量 x_k , 很自然地考虑以误差 $(x_k - \hat{y}_k(\mathbf{W}_k))$ 作为一个量测值。但是因为对 x_k 知之甚少, 只有噪声数据可以被测量, 该量测可以通过误差的方差来确定。也就是说, 要基于所有对称正定矩阵 \mathbf{W}_k 来最小化 $\text{Var}(x_k - \hat{y}_k(\mathbf{W}_k))$ 。简化记号为 $\hat{y}_k = \hat{y}_k(\mathbf{W}_k)$, 有

$$\begin{aligned} x_k - \hat{y}_k &= (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k) x_k - (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{W}_k \mathbf{z}_k \\ &= (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{W}_k (\mathbf{C}_k x_k - \mathbf{z}_k) \\ &= -(\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{W}_k \underline{\xi}_k \end{aligned}$$

根据期望运算的线性特性, 有

$$\begin{aligned}\text{Var}(\mathbf{x}_k - \hat{\mathbf{y}}_k) &= (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{W}_k E(\xi_k \xi_k^T) \mathbf{W}_k \mathbf{C}_k (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \\ &= (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{W}_k \mathbf{R}_k \mathbf{W}_k \mathbf{C}_k (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1}\end{aligned}$$

这就是将要最小化的式子。为了将该式写为完整的二次形式,需要正定对称矩阵 \mathbf{R}_k 的正平方根:设 \mathbf{R}_k 的特征值 $\lambda_1, \dots, \lambda_n$ 都为正,且 $\mathbf{R}_k = \mathbf{U}^T \text{diag}[\lambda_1, \dots, \lambda_n] \mathbf{U}$,式中 \mathbf{U} 是酉矩阵(由 λ_i ($i = 1, \dots, n$) 的正规化特征向量组成)。定义 $\mathbf{R}_k^{1/2} = \mathbf{U}^T \text{diag}[\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}] \mathbf{U}$,可以得到 $(\mathbf{R}_k^{1/2})(\mathbf{R}_k^{1/2})^T = \mathbf{R}_k$ 。则有

$$\text{Var}(\mathbf{x}_k - \hat{\mathbf{y}}_k) = \mathbf{Q}^T \mathbf{Q}$$

式中 $\mathbf{Q} = (\mathbf{R}_k^{1/2})^T \mathbf{W}_k \mathbf{C}_k (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1}$ 。

根据引理 1.1(矩阵施瓦兹不等式),在 \mathbf{P} 为 $q \times n$ 矩阵且 $\mathbf{P}^T \mathbf{P}$ 非奇异的假设下,有

$$\mathbf{Q}^T \mathbf{Q} \geq (\mathbf{P}^T \mathbf{Q})^T (\mathbf{P}^T \mathbf{P})^{-1} (\mathbf{P}^T \mathbf{Q})$$

若 $(\mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)$ 是非奇异的,令 $\mathbf{P} = (\mathbf{R}_k^{1/2})^{-1} \mathbf{C}_k$,则

$$\mathbf{P}^T \mathbf{P} = \mathbf{C}_k^T ((\mathbf{R}_k^{1/2})^T)^{-1} (\mathbf{R}_k^{1/2}) \mathbf{C}_k = \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k$$

是非奇异的,且

$$\begin{aligned}(\mathbf{P}^T \mathbf{Q})^T (\mathbf{P}^T \mathbf{P})^{-1} (\mathbf{P}^T \mathbf{Q}) &= [\mathbf{C}_k^T ((\mathbf{R}_k^{1/2})^{-1})^T (\mathbf{R}_k^{1/2})^T \mathbf{W}_k \mathbf{C}_k (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1}]^T (\mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1} \\ &\quad \cdot [\mathbf{C}_k^T ((\mathbf{R}_k^{1/2})^{-1})^T (\mathbf{R}_k^{1/2})^T \mathbf{W}_k \mathbf{C}_k (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1}] \\ &= (\mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1} \\ &= \text{Var}(\mathbf{x}_k - \hat{\mathbf{y}}_k(\mathbf{R}_k^{-1}))\end{aligned}$$

则对于所有对称正定权值矩阵 \mathbf{W}_k ,有 $\text{Var}(\mathbf{x}_k - \hat{\mathbf{y}}_k(\mathbf{W}_k)) \geq \text{Var}(\mathbf{x}_k - \hat{\mathbf{y}}_k(\mathbf{R}_k^{-1}))$ 。

因此,最优权值矩阵 $\hat{\mathbf{W}}_k = \mathbf{R}_k^{-1}$,使用该最优权值的 \mathbf{x}_k 的最优估计为

$$\hat{\mathbf{x}}_k := \hat{\mathbf{y}}_k(\mathbf{R}_k^{-1}) = (\mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{R}_k^{-1} (\mathbf{v}_k - \mathbf{D}_k \mathbf{u}_k) \quad (1.38)$$

称 $\hat{\mathbf{x}}_k$ 是状态 \mathbf{x}_k 的最小二乘最优估计。

注意到 $\hat{\mathbf{x}}_k$ 是 \mathbf{x}_k 的线性估计,是数据 $\mathbf{v}_k - \mathbf{D}_k \mathbf{u}_k$ 的线性变换,给出了 $E \hat{\mathbf{x}}_k = E \mathbf{x}_k$ 意义下 \mathbf{x}_k 的无偏估计(见练习 1.12)。对于所有正定对称权值矩阵 \mathbf{W}_k ,有

$$\text{Var}(\mathbf{x}_k - \hat{\mathbf{x}}_k) \leq \text{Var}(\mathbf{x}_k - \hat{\mathbf{y}}_k(\mathbf{W}_k))$$

即给出了 \mathbf{x}_k 的最小方差估计。

练习

1.1 证明引理 1.4。

1.2 证明引理 1.6。

1.3 举例说明存在两矩阵 \mathbf{A}, \mathbf{B} ,满足 $\mathbf{A} \geq \mathbf{B} \geq \mathbf{0}$,但是不满足不等式 $\mathbf{A}\mathbf{A}^T \geq \mathbf{B}\mathbf{B}^T$ 。

1.4 证明引理 1.8。

1.5 证明 $\int_{-\infty}^{\infty} e^{-y^2} dy = \sqrt{\pi}$ 。

1.6 证明 $\int_{-\infty}^{\infty} y^2 e^{-y^2} dy = \sqrt{\pi}/2$ 。(提示: 对积分 $-\int_{-\infty}^{\infty} e^{-xy^2} dy$, 以 x 为变量取微分, 然后令 $x \rightarrow 1$)

1.7 如果 $f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} (\det \mathbf{R})^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \underline{\mu})^T \mathbf{R}^{-1} (\mathbf{x} - \underline{\mu}) \right\}$,

证明:

(a)

$$\begin{aligned} E(\mathbf{X}) &= \int_{-\infty}^{+\infty} \mathbf{x} f(\mathbf{x}) d\mathbf{x} \\ &:= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} f(\mathbf{x}) dx_1 \cdots dx_n \\ &= \underline{\mu} \end{aligned}$$

(b)

$$\text{Var}(\mathbf{X}) = E(\mathbf{X} - \underline{\mu})(\mathbf{X} - \underline{\mu})^T = \mathbf{R}$$

1.8 证明式(1.32)关于期望、方差和协方差的性质。

1.9 证明服从正态分布的随机变量 $\mathbf{X}_1, \mathbf{X}_2$, 当且仅当 $\text{Cov}(\mathbf{X}_1, \mathbf{X}_2) = 0$ 时, $\mathbf{X}_1, \mathbf{X}_2$ 是独立的。

1.10 证明式(1.35)。

1.11 考虑对所有的 n 维向量 \mathbf{y}_k 最小化下面的量: $F(\mathbf{y}_k) = E(\mathbf{z}_k - \mathbf{C}_k \mathbf{y}_k)^T \mathbf{W}_k (\mathbf{z}_k - \mathbf{C}_k \mathbf{y}_k)$, 式中, \mathbf{z}_k 是一个 $q \times 1$ 的向量, \mathbf{C}_k 是 $q \times n$ 的矩阵, \mathbf{W}_k 是 $q \times q$ 的权值矩阵, 那么矩阵 $(\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)$ 是非奇异的。通过令 $dF(\mathbf{y}_k)/d\mathbf{y}_k = \mathbf{0}$, 证明最优解 $\hat{\mathbf{y}}_k$ 由下式给出: $\hat{\mathbf{y}}_k = (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{W}_k \mathbf{z}_k$ 。

(提示: 标量函数 $F(\mathbf{y})$ 对 n 维向量 $\mathbf{y} = [y_1 \cdots y_n]^T$ 的微分定义为 $\frac{dF(\mathbf{y})}{d\mathbf{y}} =$

$$\left[\frac{\partial F}{\partial y_1} \cdots \frac{\partial F}{\partial y_n} \right]^T$$

1.12 证明式(1.38)给出的估计 \hat{x}_k 在 $E\hat{x}_k = E\mathbf{x}_k$ 的条件下是 \mathbf{x}_k 的无偏估计。

第 2 章

卡尔曼滤波：简单推导

这一章给出卡尔曼滤波算法的最基本介绍。应用所获取的数据，基于状态向量的最小二乘无偏最优估计和最优权值，通过假设某些矩阵是可逆的，推导出卡尔曼滤波“预测-校正”算法。滤波算法首先针对无确定性(控制)输入系统求解，再通过叠加确定性解，就可以得到常规卡尔曼滤波算法。

2.1 模型

考虑在状态空间描述的线性系统：

$$\begin{cases} \mathbf{y}_{k+1} = \mathbf{A}_k \mathbf{y}_k + \mathbf{B}_k \mathbf{u}_k + \boldsymbol{\Gamma}_k \underline{\boldsymbol{\xi}}_k \\ \mathbf{w}_k = \mathbf{C}_k \mathbf{y}_k + \mathbf{D}_k \mathbf{u}_k + \underline{\boldsymbol{\eta}}_k \end{cases}$$

式中 $\mathbf{A}_k, \mathbf{B}_k, \boldsymbol{\Gamma}_k, \mathbf{C}_k, \mathbf{D}_k$ 分别是 $n \times n, n \times m, n \times p, q \times n, q \times m$ 阶常值矩阵(已知)，且有 $1 \leq m, p, q \leq n$ ， $\{\mathbf{u}_k\}$ 是 m 维向量序列(称为确定性输入序列)(已知)， $\{\underline{\boldsymbol{\xi}}_k\}$ 和 $\{\underline{\boldsymbol{\eta}}_k\}$ 分别是已知均值、方差和协方差等统计信息的系统和观测噪声序列(未知)。因为同时有确定性输入 $\{\mathbf{u}_k\}$ 和噪声序列 $\{\underline{\boldsymbol{\xi}}_k\}$ 及 $\{\underline{\boldsymbol{\eta}}_k\}$ ，该系统被称为线性确定性/随机系统。该系统能够分解成一个线性确定性系统：

$$\begin{cases} \mathbf{z}_{k+1} = \mathbf{A}_k \mathbf{z}_k + \mathbf{B}_k \mathbf{u}_k \\ \mathbf{s}_k = \mathbf{C}_k \mathbf{z}_k + \mathbf{D}_k \mathbf{u}_k \end{cases}$$

与一个线性(完全)随机系统的和

$$\begin{cases} \underline{x}_{k+1} = \mathbf{A}_k \underline{x}_k + \Gamma_k \underline{\xi}_k \\ \underline{v}_k = \mathbf{C}_k \underline{x}_k + \underline{\eta}_k \end{cases} \quad (2.1)$$

式中 $w_k = s_k + v_k$, $y_k = z_k + x_k$ 。这样分解的好处是线性确定性系统的解 z_k 能够由转换方程给出

$$z_k = (\mathbf{A}_{k-1} \cdots \mathbf{A}_0) z_0 + \sum_{i=1}^k (\mathbf{A}_{k-1} \cdots \mathbf{A}_{i-1}) \mathbf{B}_{i-1} u_{i-1}$$

而通过求解式(2.1)描述的随机状态空间的 x_k 的最优估计 \hat{x}_k , 从而

$$\hat{y}_k = z_k + \hat{x}_k$$

为原始线性系统的状态向量 y_k 的最优估计。当然, 估计必须依赖于噪声序列的统计信息。在本章中, 只考虑零均值高斯白噪声过程。

假设 2.1 令 $\{\underline{\xi}_k\}$ 和 $\{\underline{\eta}_k\}$ 是零均值高斯白噪声序列。那么对于所有 k 和 l , $\text{Var}(\underline{\xi}_k) = Q_k$ 和 $\text{Var}(\underline{\eta}_k) = R_k$ 是正定矩阵, 且 $E(\underline{\xi}_k \underline{\eta}_l^\top) = \mathbf{0}$ 。另假设初始状态 x_0 与 $\underline{\xi}_k$ 、 $\underline{\eta}_k$ 独立, 即对于所有的 k , 有 $E(x_0 \underline{\xi}_k^\top) = \mathbf{0}$ 和 $E(x_0 \underline{\eta}_k^\top) = \mathbf{0}$ 成立。 ■■■

2.2 最优准则

由 1.3 可知, 在确定 x_k 的最优估计 \hat{x}_k 时, 最优性是通过选择最优权值矩阵给出的最小二乘意义下的最小方差估计取得的。但是需要联合所有数据 v_j ($j = 0, 1, \dots, k$) 的信息来确定 x_k 的估计 \hat{x}_k (而不是 1.3 节讨论的仅仅利用 v_k)。为了实现该思路, 引入向量

$$\bar{v}_j = \begin{bmatrix} v_0 \\ \vdots \\ v_j \end{bmatrix}, \quad j = 0, 1, \dots$$

并从数据向量 \bar{v}_k 中求得 \hat{x}_k 。为了完成该过程, 假设到当前时刻的所有系统矩阵 A_j 非奇异。那么状态空间描述的线性随机系统可以写为

$$\bar{v}_j = H_{k,j} x_k + \bar{\epsilon}_{k,j} \quad (2.2)$$

式中:

$$H_{k,j} = \begin{bmatrix} C_0 \Phi_{0k} \\ \vdots \\ C_j \Phi_{jk} \end{bmatrix} \quad \text{和} \quad \bar{\epsilon}_{k,j} = \begin{bmatrix} \epsilon_{k,0} \\ \vdots \\ \epsilon_{k,j} \end{bmatrix}$$

转移矩阵 Φ_{lk} 定义为

$$\Phi_{lk} = \begin{cases} A_{l-1} \cdots A_k, & l > k \\ I, & l = k \end{cases}$$

当 $l < k$ 时, $\Phi_{lk} = \Phi_{kl}^{-1}$,且

$$\underline{\epsilon}_{k,l} = \underline{\eta}_l - C_l \sum_{i=l+1}^k \Phi_{li} \Gamma_{i-1} \underline{\xi}_{i-1}$$

应用前面介绍的 Φ_{lk} 的逆变换特性,转移方程为

$$\underline{x}_k = \Phi_{kl} \underline{x}_l + \sum_{i=l+1}^k \Phi_{ki} \Gamma_{i-1} \underline{\xi}_{i-1}$$

该式可以轻易地从式(2.1)介绍的系统方程得到,有

$$\underline{x}_l = \Phi_{lk} \underline{x}_k - \sum_{i=l+1}^k \Phi_{li} \Gamma_{i-1} \underline{\xi}_{i-1}$$

可得

$$\begin{aligned} H_{k,j} \underline{x}_k + \bar{\epsilon}_{k,j} &= \begin{bmatrix} C_0 & \Phi_{0k} \\ \vdots & \\ C_j & \Phi_{jk} \end{bmatrix} \underline{x}_k + \begin{bmatrix} \underline{\eta}_0 - C_0 \sum_{i=1}^k \Phi_{0i} \Gamma_{i-1} \underline{\xi}_{i-1} \\ \vdots \\ \underline{\eta}_j - C_j \sum_{i=j+1}^k \Phi_{ji} \Gamma_{i-1} \underline{\xi}_{i-1} \end{bmatrix} \\ &= \begin{bmatrix} C_0 \underline{x}_0 + \underline{\eta}_0 \\ \vdots \\ C_j \underline{x}_j + \underline{\eta}_j \end{bmatrix} = \begin{bmatrix} v_0 \\ \vdots \\ v_j \end{bmatrix} = \bar{v}_j \end{aligned}$$

即式(2.2)。

应用1.3节讨论的最小二乘估计,权值为 $W_{k,j} = (\text{Var}(\bar{\epsilon}_{k,j}))^{-1}$,这样通过使用数据 v_0, \dots, v_j ,就可以得到 \underline{x}_k 的线性、无偏、最小方差最小二乘估计 $\hat{x}_{k|j}$ 。

定义2.1 (1)对于 $j=k$,定义 $\hat{x}_k = \hat{x}_{k|k}$,并称该估计过程为数字滤波过程;(2)对于 $j < k$,定义 $\hat{x}_{k|j}$ 为 \underline{x}_k 的最优预测,并称该过程为数字预测过程;(3)对于 $j > k$,定义 $\hat{x}_{k|j}$ 为 \underline{x}_k 的平滑估计,并且称该过程为数字平滑过程。■

本书只讨论数字滤波。由于 $\hat{x}_k = \hat{x}_{k|k}$ 是根据所有数据 v_0, \dots, v_j 确定的,因为数据存储量和计算量随着时间增加,该方法不适用于 k 值很大的实时问题。因此我们打算推导从“预测” $\hat{x}_{k|k-1}$ 得到 $\hat{x}_k = \hat{x}_{k|k}$,及从估计 $\hat{x}_{k-1} = \hat{x}_{k-1|k-1}$ 得到 $\hat{x}_{k|k-1}$ 的递推公式。在其中的每一步,由于只使用最新的数据信息,故只需用很小的数据存储量。这就是通常提到的卡尔曼滤波算法。

2.3 预测-校正公式

为了实时计算 \hat{x}_k , 本节将推导递推公式:

$$\begin{cases} \hat{x}_{k|k} = \hat{x}_{k|k-1} + G_k(v_k - C_k \hat{x}_{k|k-1}) \\ \hat{x}_{k|k-1} = A_{k-1} \hat{x}_{k-1|k-1} \end{cases}$$

式中 G_k 为卡尔曼增益矩阵。

开始点是初始估计 $\hat{x}_0 = \hat{x}_{0|0}$, 因为 \hat{x}_0 是初始状态 x_0 的无偏估计, 可以使用常值向量 $\hat{x}_0 = E(x_0)$ 。而在实际的卡尔曼滤波中, G_k 也必须递推计算。这两个递推过程合起来称为卡尔曼滤波过程。

选择权值矩阵:

$$W_{k,j} = (\text{Var}(\underline{\epsilon}_{k,j}))^{-1}$$

使用式(2.2)的 \bar{v}_j (见 1.3 节), 使得 $\hat{x}_{k|j}$ 是 x_k 的具有最小方差的(最优)最小二乘估计。易证:

$$W_{k,k-1}^{-1} = \begin{bmatrix} R_0 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & R_{k-1} \end{bmatrix} + \text{Var} \begin{bmatrix} C_0 \sum_{i=1}^k \Phi_{0i} \Gamma_{i-1} \xi_{i-1} \\ \vdots \\ C_{k-1} \Phi_{k-1,k} \Gamma_{k-1} \xi_{k-1} \end{bmatrix} \quad (2.3)$$

$$W_{k,k}^{-1} = \begin{bmatrix} W_{k,k-1}^{-1} & \mathbf{0} \\ \mathbf{0} & R_k \end{bmatrix} \quad (2.4)$$

(见练习 2.1)。所以, $W_{k,k-1}$ 和 $W_{k,k}$ 是正定的(见练习 2.2)。

在本章中, 假设矩阵 $(H_{k,j}^T W_{k,j} H_{k,j})$, $j=k-1, k$, 非奇异。^①

由 1.3 节可知:

$$\hat{x}_{k|j} = (H_{k,j}^T W_{k,j} H_{k,j})^{-1} H_{k,j}^T W_{k,j} \bar{v}_j \quad (2.5)$$

我们第一个目标是建立 $\hat{x}_{k|k-1}$ 与 $\hat{x}_{k|k}$ 的联系。为了实现该目标, 注意到:

$$\begin{aligned} H_{k,k}^T W_{k,k} H_{k,k} &= [H_{k,k-1}^T C_k^T] \begin{bmatrix} W_{k,k-1} & \mathbf{0} \\ \mathbf{0} & R_k^{-1} \end{bmatrix} [H_{k,k-1} \\ &\quad C_k] \\ &= H_{k,k-1}^T W_{k,k-1} H_{k,k-1} + C_k^T R_k^{-1} C_k \end{aligned}$$

及

^① 译者注: 以后将除去这一假定。

$$\mathbf{H}_{k,k}^T \mathbf{W}_{k,k} \bar{\mathbf{v}}_k = \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \bar{\mathbf{v}}_{k-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{v}_k$$

应用式(2.5)和前面的两个方程,得

$$\begin{aligned} (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k) \hat{\mathbf{x}}_{k|k-1} &= \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \bar{\mathbf{v}}_{k-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1} \\ (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k) \hat{\mathbf{x}}_{k|k} &= (\mathbf{H}_{k,k}^T \mathbf{W}_{k,k} \mathbf{H}_{k,k}) \hat{\mathbf{x}}_{k|k} \\ &= \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \bar{\mathbf{v}}_{k-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{v}_k \end{aligned}$$

通过简单的减法可得

$$(\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k) (\hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k-1}) = \mathbf{C}_k^T \mathbf{R}_k^{-1} (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})$$

定义

$$\begin{aligned} \mathbf{G}_k &= (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{R}_k^{-1} \\ &= (\mathbf{H}_{k,k}^T \mathbf{W}_{k,k} \mathbf{H}_{k,k})^{-1} \mathbf{C}_k^T \mathbf{R}_k^{-1} \end{aligned}$$

这样就得到

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \quad (2.6)$$

因为 $\hat{\mathbf{x}}_{k|k-1}$ 是一步预测, $(\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})$ 是实际数据和预测之间的误差,式(2.6)实际上是以卡尔曼滤波增益 \mathbf{G}_k 作为权值矩阵的“预测-校正”公式。为了完成递推过程,还需要一个从 $\hat{\mathbf{x}}_{k-1|k-1}$ 到 $\hat{\mathbf{x}}_{k|k-1}$ 的公式

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} \quad (2.7)$$

为了证明该式,首先注意到

$$\underline{\boldsymbol{\epsilon}}_{k,k-1} = \underline{\boldsymbol{\epsilon}}_{k-1,k-1} - \mathbf{H}_{k,k-1} \Gamma_{k-1} \underline{\boldsymbol{\xi}}_{k-1}$$

使得

$$\mathbf{W}_{k,k-1}^{-1} = \mathbf{W}_{k-1,k-1}^{-1} + \mathbf{H}_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} \mathbf{Q}_{k-1} \Gamma_{k-1}^T \Phi_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \quad (2.8)$$

(见练习2.3)。

根据引理1.2,有

$$\begin{aligned} \mathbf{W}_{k,k-1} &= \mathbf{W}_{k-1,k-1} - \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} (\mathbf{Q}_{k-1}^{-1} \\ &\quad + \Gamma_{k-1}^T \Phi_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\ &\quad \cdot \Gamma_{k-1}^T \Phi_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \end{aligned} \quad (2.9)$$

(见练习2.4)。

然后根据转换关系

$$\mathbf{H}_{k,k-1} = \mathbf{H}_{k-1,k-1} \Phi_{k-1,k}$$

有

$$\begin{aligned} \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} &= \Phi_{k-1,k}^T \{ \mathbf{I} - \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} (\mathbf{Q}_{k-1}^{-1} \\ &\quad + \Gamma_{k-1}^T \Phi_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\ &\quad \cdot \Gamma_{k-1}^T \Phi_{k-1,k}^T \} \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \end{aligned} \quad (2.10)$$

(见练习 2.5)。则

$$\begin{aligned} & (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1}) \Phi_{k,k-1} (\mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1})^{-1} \\ & \cdot \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} = \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \end{aligned} \quad (2.11)$$

(见练习 2.6)。结合式(2.5), 当 $j=k-1$ 和 k 时得到式(2.7)。

下一个目标是得到卡尔曼增益矩阵 \mathbf{G}_k 的递推公式。首先有

$$\mathbf{G}_k = \mathbf{P}_{k,k} \mathbf{C}_k^T \mathbf{R}_k^{-1}$$

式中

$$\mathbf{P}_{k,k} = (\mathbf{H}_{k,k}^T \mathbf{W}_{k,k} \mathbf{H}_{k,k})^{-1}$$

且令

$$\mathbf{P}_{k,k-1} = (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1}$$

又因

$$\mathbf{P}_{k,k}^{-1} = \mathbf{P}_{k,k-1}^{-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k$$

应用引理 1.2, 可得

$$\mathbf{P}_{k,k} = \mathbf{P}_{k,k-1} - \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \mathbf{C}_k \mathbf{P}_{k,k-1}$$

可以证明：

$$\mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \quad (2.12)$$

(见练习 2.7)。因此

$$\mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \quad (2.13)$$

此外, 还有

$$\mathbf{P}_{k,k-1} = \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T \quad (2.14)$$

(见练习 2.8)。

应用式(2.13)和式(2.14)及初始矩阵 $\mathbf{P}_{0,0}$, 可得 $\mathbf{P}_{k-1,k-1}, \mathbf{P}_{k,k-1}, \mathbf{G}_k$ 和 $\mathbf{P}_{k,k} (k=1, 2, \dots)$ 的递推计算方法。首先有

$$\begin{aligned} \mathbf{P}_{k,k-1} &= E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T \\ &= \text{Var}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \end{aligned} \quad (2.15)$$

(见练习 2.9), 还有

$$\mathbf{P}_{k,k} = E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T = \text{Var}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}) \quad (2.16)$$

特别地, 当 $k=0$ 时, 有

$$\mathbf{P}_{0,0} = E(\mathbf{x}_0 - E\mathbf{x}_0)(\mathbf{x}_0 - E\mathbf{x}_0)^T = \text{Var}(\mathbf{x}_0)$$

最后, 联合上面得到的所有结果, 得到式(2.1)所示的状态空间描述的线性随机系统的卡尔曼滤波过程：

$$\left\{ \begin{array}{l} \mathbf{P}_{0,0} = \text{Var}(\mathbf{x}_0) \\ \mathbf{P}_{k,k-1} = \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T \\ \mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ \mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \\ \hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \\ \hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \\ k = 1, 2, \dots \end{array} \right. \quad (2.17)$$

该算法可以通过图 2.1 的方式来实现。

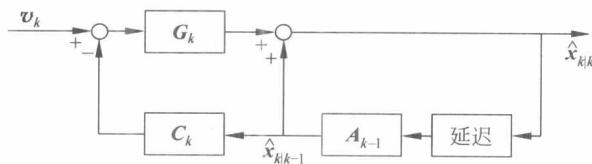


图 2.1 卡尔曼滤波算法结构框图

2.4 卡尔曼滤波过程

现在考虑具有确定性控制输入 $\{u_k\}$ 的常规线性确定性/随机系统。考虑状态空间模型

$$\left\{ \begin{array}{l} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k u_k + \boldsymbol{\Gamma}_k \xi_k \\ \mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k u_k + \eta_k \end{array} \right.$$

式中 $\{u_k\}$ 是 m 维向量序列 ($1 \leq m \leq n$)。

将确定性解叠加到式(2.17)上，则可得到该系统的卡尔曼滤波过程：

$$\left\{ \begin{array}{l} \mathbf{P}_{0,0} = \text{Var}(\mathbf{x}_0) \\ \mathbf{P}_{k,k-1} = \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T \\ \mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ \mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \\ \hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \\ \hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_{k-1} u_{k-1} \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1} - \mathbf{D}_k u_k) \\ k = 1, 2, \dots \end{array} \right. \quad (2.18)$$

(见练习 2.13)。该算法可以通过图 2.2 的方式来实现。

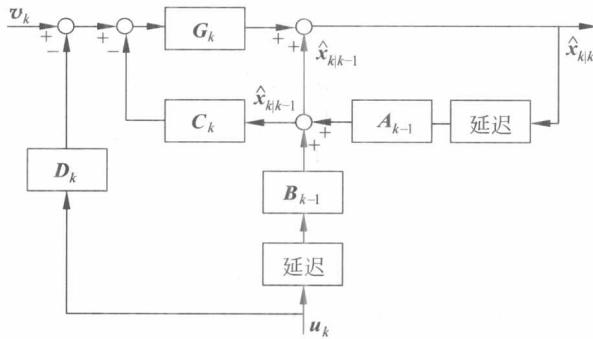


图 2.2 有输入的卡尔曼滤波算法结构框图

练习

2.1 令

$$\underline{\epsilon}_{k,j} = \begin{bmatrix} \underline{\epsilon}_{k,0} \\ \vdots \\ \underline{\epsilon}_{k,j} \end{bmatrix} \quad \text{和} \quad \underline{\epsilon}_{k,l} = \underline{\eta}_l - C_l \sum_{i=l+1}^k \Phi_{li} \Gamma_{i-1} \underline{\xi}_{i-1}$$

式中 $\{\underline{\xi}_k\}$ 和 $\{\underline{\eta}_k\}$ 都是零均值高斯白噪声序列, 且 $\text{Var}(\underline{\xi}_k) = Q_k$ 和 $\text{Var}(\underline{\eta}_k) = R_k$ 。
定义 $W_{k,j} = (\text{Var}(\underline{\epsilon}_{k,j}))^{-1}$ 。证明:

$$W_{k,k-1}^{-1} = \begin{bmatrix} R_0 & & 0 \\ & \ddots & \\ 0 & & R_{k-1} \end{bmatrix} + \text{Var} \begin{bmatrix} C_0 \sum_{i=1}^k \Phi_{0i} \Gamma_{i-1} \underline{\xi}_{i-1} \\ \vdots \\ C_{k-1} \Phi_{k-1,k} \Gamma_{k-1} \underline{\xi}_{k-1} \end{bmatrix}$$

和

$$W_{k,k}^{-1} = \begin{bmatrix} W_{k,k-1}^{-1} & 0 \\ 0 & R_k \end{bmatrix}$$

2.2 证明正定矩阵 A 和一个非负定矩阵 B 的和是正定的。

2.3 设 $\underline{\epsilon}_{k,j}$ 和 $W_{k,j}$ 与练习 2.1 的定义一样。证明下列关系: $\underline{\epsilon}_{k,k-1} = \underline{\epsilon}_{k-1,k-1} -$

$\mathbf{H}_{k,k-1} \boldsymbol{\Gamma}_{k-1} \boldsymbol{\xi}_{k-1}$, 式中, $\mathbf{H}_{k,j} = \begin{bmatrix} \mathbf{C}_0 \boldsymbol{\Phi}_{0k} \\ \vdots \\ \mathbf{C}_j \boldsymbol{\Phi}_{jk} \end{bmatrix}$, 然后再证明:

$$\mathbf{W}_{k,k-1}^{-1} = \mathbf{W}_{k-1,k-1}^{-1} + \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T \boldsymbol{\Phi}_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T$$

2.4 应用练习 2.3 和引理 1.2, 证明:

$$\begin{aligned} \mathbf{W}_{k,k-1} &= \mathbf{W}_{k-1,k-1} - \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1} (\mathbf{Q}_{k-1}^{-1} \\ &\quad + \boldsymbol{\Gamma}_{k-1}^T \boldsymbol{\Phi}_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1})^{-1} \\ &\quad \cdot \boldsymbol{\Gamma}_{k-1}^T \boldsymbol{\Phi}_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \end{aligned}$$

2.5 应用练习 2.4 和 $\mathbf{H}_{k,k-1} = \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k}$, 证明:

$$\begin{aligned} \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} &= \boldsymbol{\Phi}_{k-1,k}^T \{ \mathbf{I} - \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1} (\mathbf{Q}_{k-1}^{-1} \\ &\quad + \boldsymbol{\Gamma}_{k-1}^T \boldsymbol{\Phi}_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1})^{-1} \\ &\quad \cdot \boldsymbol{\Gamma}_{k-1}^T \boldsymbol{\Phi}_{k-1,k}^T \} \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \end{aligned}$$

2.6 应用练习 2.5, 推导下列等式:

$$\begin{aligned} (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1}) \boldsymbol{\Phi}_{k,k-1} (\mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1})^{-1} \\ \cdot \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} = \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \end{aligned}$$

2.7 应用引理 1.2, 证明:

$$\mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} = \mathbf{P}_{k,k} \mathbf{C}_k^T \mathbf{R}_k^{-1} = \mathbf{G}_k$$

2.8 从 $\mathbf{P}_{k,k-1} = (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1}$ 开始, 应用引理 1.2、式(2.8)及定义 $\mathbf{P}_{k,k} = (\mathbf{H}_{k,k}^T \mathbf{W}_{k,k} \mathbf{H}_{k,k})^{-1}$, 证明:

$$\mathbf{P}_{k,k-1} = \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T$$

2.9 应用式(2.5)和式(2.2), 证明:

$$E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T = \mathbf{P}_{k,k-1}$$

和

$$E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T = \mathbf{P}_{k,k}$$

2.10 考虑一维线性随机动态系统:

$$x_{k+1} = ax_k + \xi_k, \quad x_0 = 0$$

式中 $E(x_k) = 0$, $\text{Var}(x_k) = \sigma^2$, $E(x_k \xi_j) = 0$, $E(\xi_k) = 0$, $E(\xi_k \xi_j) = \mu^2 \delta_{kj}$ 。证明对于所有整数 j , 有 $\sigma^2 = \mu^2 / (1 - a^2)$, $E(x_k x_{k+j}) = a^{|j|} \sigma^2$ 。

2.11 考虑一维随机线性系统:

$$\begin{cases} x_{k+1} = x_k \\ v_k = x_k + \eta_k \end{cases}$$

式中 $E(\eta_k) = 0$, $\text{Var}(\eta_k) = \sigma^2$, $E(x_0) = 0$, $\text{Var}(x_0) = \mu^2$ 。证明:

$$\begin{cases} \hat{x}_{k|k} = \hat{x}_{k-1|k-1} + \frac{\mu^2}{\sigma^2 + k\mu^2} (v_k - \hat{x}_{k-1|k-1}) \\ \hat{x}_{0|0} = 0 \end{cases}$$

且当 $k \rightarrow \infty$ 时, $\hat{x}_{k|k} \rightarrow c$ (c 为一常值)。

2.12 令 $\{v_k\}$ 是对一个方差 Q 未知的零均值随机向量 y 的观测序列。 y 的方差可以通过下式估计: $\hat{Q}_N = \frac{1}{N} \sum_{k=1}^N (v_k v_k^T)$ 。试为该估计过程推导一个预测-校正递推公式。

2.13 考虑线性确定性/随机系统:

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k + \Gamma_k \xi_k \\ v_k = C_k x_k + D_k u_k + \eta_k \end{cases}$$

式中 $\{u_k\}$ 是给定的 m 维 ($1 \leq m \leq n$) 确定性控制输入, 满足假设 2.1 且矩阵 $\text{Var}(\xi_{k,j})$ 非奇异 ($\xi_{k,j}$ 的详细定义见式(2.2))。推导该模型的卡尔曼滤波方程。

2.14 在数字信号处理中, 一个被广泛使用的数学模型是自回归滑动平均 (autoregressive moving-average, ARMA) 过程:

$$v_k = \sum_{i=1}^N B_i v_{k-i} + \sum_{i=0}^M A_i u_{k-i}$$

式中 $n \times n$ 阶矩阵 B_1, \dots, B_N 和 $n \times q$ 阶矩阵 A_0, A_1, \dots, A_M 关于时间变量 k 是独立的, 且 $\{u_k\}$ 和 $\{v_k\}$ 分别是输入和输出数字信号序列 (见图 2.3)。假设 $M \leq N$, 证明输入-输出关系可以用下列状态空间模型来描述:

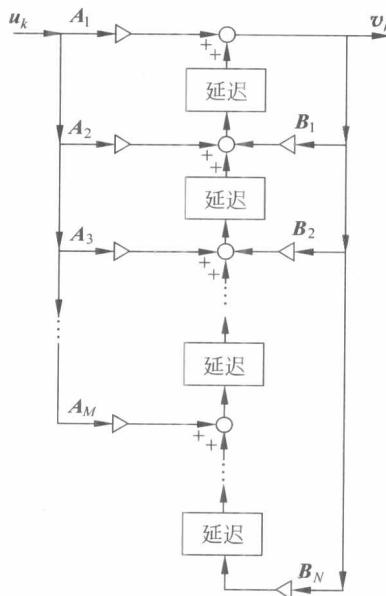


图 2.3 ARMA 模型的输入输出序列

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \\ \mathbf{v}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k \end{cases}$$

其中 $\mathbf{x}_0 = \mathbf{0}$,

$$\mathbf{A} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{B}_2 & \mathbf{0} & \mathbf{I} & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \mathbf{B}_{N-1} & \mathbf{0} & \cdots & \cdots & \mathbf{I} \\ \mathbf{B}_N & \mathbf{0} & \cdots & \cdots & \mathbf{0} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{A}_1 + \mathbf{B}_1\mathbf{A}_0 \\ \mathbf{A}_2 + \mathbf{B}_2\mathbf{A}_0 \\ \vdots \\ \mathbf{A}_M + \mathbf{B}_M\mathbf{A}_0 \\ \mathbf{B}_{M+1}\mathbf{A}_0 \\ \vdots \\ \mathbf{B}_N\mathbf{A}_0 \end{bmatrix}$$

$$\mathbf{C} = [\mathbf{I} \quad \mathbf{0} \quad \cdots \quad \mathbf{0}] \quad \text{和} \quad \mathbf{D} = [\mathbf{A}_0]$$

第 3 章

正交投影和卡尔曼滤波

第 2 章讨论了最优卡尔曼滤波的基本推导过程, 其优点是状态向量 x_k 的最优估计 $\hat{x}_k = \hat{x}_{k|k}$ 可以很容易理解为 x_k 的最小二乘估计。它具有如下的特性: (1) 从数据 $\bar{v}_k = [v_0^T \quad \cdots \quad v_k^T]^T$ 得到 \hat{x}_k 的变换是线性的; (2) 在 $E \hat{x}_k = E x_k$ 意义下, \hat{x}_k 是无偏的; (3) 以 $(\text{Var}(\underline{\epsilon}_{k,k}))^{-1}$ 为最优权值, 得到了最小方差估计。其缺点是, 必须对一些矩阵作非奇异的假设。在本章中, 我们将放弃非奇异假设(见式(2.4)), 给出卡尔曼滤波算法的严格推导。

3.1 最优估计的正交性

考虑满足假设 2.1 的由式(2.1)描述的线性随机系统, 即状态空间模型:

$$\begin{cases} x_{k+1} = A_k x_k + \Gamma_k \xi_k \\ v_k = C_k x_k + \eta_k \end{cases} \quad (3.1)$$

式中 A_k 、 Γ_k 和 C_k 分别为已知的 $n \times n$ 、 $n \times p$ 和 $q \times n$ ($1 \leqslant p, q \leqslant n$) 常值矩阵, 且对于所有 $k, l = 0, 1, \dots$, 有

$$E(\xi_k) = \mathbf{0}, \quad E(\xi_k \xi_l^T) = Q_k \delta_{kl}$$

$$E(\eta_k) = \mathbf{0}, \quad E(\eta_k \eta_l^T) = R_k \delta_{kl}$$

$$E(\xi_k \eta_l^T) = \mathbf{0}, \quad E(x_0 \xi_k^T) = \mathbf{0}, \quad E(x_0 \eta_k^T) = \mathbf{0}$$

式中 Q_k 和 R_k 为正定对称矩阵。

设 \mathbf{x} 是 n 维随机向量, \mathbf{w} 是 q 维随机向量。定义“内积” $\langle \mathbf{x}, \mathbf{w} \rangle$ 是 $n \times q$ 阶矩阵:

$$\langle \mathbf{x}, \mathbf{w} \rangle = \text{Cov}(\mathbf{x}, \mathbf{w}) = E[(\mathbf{x} - E(\mathbf{x}))(\mathbf{w} - E(\mathbf{w}))^\top]$$

设 $\|\mathbf{w}\|_q$ 是 $\langle \mathbf{w}, \mathbf{w} \rangle$ 的正平方根, 即 $\|\mathbf{w}\|_q$ 是一个非负定 $q \times q$ 阶矩阵, 且

$$\|\mathbf{w}\|_q^2 = \|\mathbf{w}\|_q \|\mathbf{w}\|_q^\top = \langle \mathbf{w}, \mathbf{w} \rangle$$

同理, 设 $\|\mathbf{x}\|_n$ 是 $\langle \mathbf{x}, \mathbf{x} \rangle$ 的正平方根, $\mathbf{w}_0, \dots, \mathbf{w}_r$ 是 q 维随机向量, 并考虑线性生成空间:

$$Y(\mathbf{w}_0, \dots, \mathbf{w}_r) = \left\{ \mathbf{y}: \mathbf{y} = \sum_{i=0}^r \mathbf{P}_i \mathbf{w}_i, \quad \mathbf{P}_0, \dots, \mathbf{P}_r \text{ 是 } n \times q \text{ 常值矩阵} \right\}$$

需要研究的第一个最小化问题是在 $Y(\mathbf{w}_0, \dots, \mathbf{w}_r)$ 中确定一个 $\hat{\mathbf{y}}$, 使得 $\text{tr} \|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k$, 其中:

$$F_k := \min \{ \text{tr} \|\mathbf{x}_k - \mathbf{y}\|_n^2 : \mathbf{y} \in Y(\mathbf{w}_0, \dots, \mathbf{w}_r) \} \quad (3.2)$$

下面给出 $\hat{\mathbf{y}}$ 的特征。

引理 3.1 $\hat{\mathbf{y}} \in Y(\mathbf{w}_0, \dots, \mathbf{w}_r)$ 满足 $\text{tr} \|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k$ 当且仅当

$$\langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_j \rangle = \mathbf{O}_{n \times q}$$

对于所有 $j=0, 1, \dots, r$ 成立。 $\hat{\mathbf{y}}$ 在下式的意义下是唯一的: 只有当 $\hat{\mathbf{y}} = \bar{\mathbf{y}}$ 时,

$$\text{tr} \|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = \text{tr} \|\mathbf{x}_k - \bar{\mathbf{y}}\|_n^2$$

证明如下。

首先假设 $\text{tr} \|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k$, 但是对于满足 $0 \leq j_0 \leq r$ 的 j_0 , 有 $\langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_{j_0} \rangle = \mathbf{C} \neq \mathbf{O}_{n \times q}$ 。

因为 $\mathbf{w}_{j_0} \neq 0$, 故 $\|\mathbf{w}_{j_0}\|_q^2$ 是一个正定对称矩阵, 其逆 $\|\mathbf{w}_{j_0}\|_q^{-2}$ 也是对称正定的。因此, $\mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2} \mathbf{C}^\top \neq \mathbf{O}_{n \times n}$, 且为一个非负定对称矩阵。可以证明:

$$\text{tr}\{\mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2} \mathbf{C}^\top\} > 0 \quad (3.3)$$

(见练习 3.1)。

向量 $\hat{\mathbf{y}} + \mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2} \mathbf{w}_{j_0}$ 在 $Y(\mathbf{w}_0, \dots, \mathbf{w}_r)$ 中, 并且根据式(3.3)有

$$\begin{aligned} & \text{tr} \|\mathbf{x}_k - (\hat{\mathbf{y}} + \mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2} \mathbf{w}_{j_0})\|_n^2 \\ &= \text{tr} \{ \|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 - \langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_{j_0} \rangle (\mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2})^\top - \mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2} \langle \mathbf{w}_{j_0}, \mathbf{x}_k - \hat{\mathbf{y}} \rangle \\ & \quad + \mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2} \|\mathbf{w}_{j_0}\|_q^2 (\mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2})^\top \} \\ &= \text{tr} \{ \|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 - \mathbf{C} \|\mathbf{w}_{j_0}\|_q^{-2} \mathbf{C}^\top \} \\ &< \text{tr} \|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k \end{aligned}$$

这与式(3.2)中 F_k 的定义相矛盾。

相反地,对于所有的 $j=0,1,\dots,r$,设 $\langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_j \rangle = \mathbf{0}_{n \times q}$ 。

令 \mathbf{y} 是 $Y(\mathbf{w}_0, \dots, \mathbf{w}_r)$ 中一个任意的 n 维向量,写为 $\mathbf{y}_0 = \mathbf{y} - \hat{\mathbf{y}} = \sum_{j=1}^r \mathbf{P}_{0j} \mathbf{w}_j$,

其中 $\mathbf{P}_{0j} (j=0,1,\dots,r)$ 是 $n \times q$ 阶常值矩阵,则

$$\begin{aligned} & \text{tr} \| \mathbf{x}_k - \mathbf{y} \|_n^2 \\ &= \text{tr} \| (\mathbf{x}_k - \hat{\mathbf{y}}) - \mathbf{y}_0 \|_n^2 \\ &= \text{tr} \{ \| \mathbf{x}_k - \hat{\mathbf{y}} \|_n^2 - \langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{y}_0 \rangle - \langle \mathbf{y}_0, \mathbf{x}_k - \hat{\mathbf{y}} \rangle + \| \mathbf{y}_0 \|_n^2 \} \\ &= \text{tr} \left\{ \| \mathbf{x}_k - \hat{\mathbf{y}} \|_n^2 - \sum_{j=0}^r \langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_j \rangle \mathbf{P}_{0j}^T - \sum_{j=0}^r \mathbf{P}_{0j} \langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_j \rangle^T + \| \mathbf{y}_0 \|_n^2 \right\} \\ &= \text{tr} \| \mathbf{x}_k - \hat{\mathbf{y}} \|_n^2 + \text{tr} \| \mathbf{y}_0 \|_n^2 \\ &\geq \text{tr} \| \mathbf{x}_k - \hat{\mathbf{y}} \|_n^2 \end{aligned}$$

所以 $\text{tr} \| \mathbf{x}_k - \hat{\mathbf{y}} \|_n^2 = F_k$ 。更进一步,等号可以达到当且仅当 $\text{tr} \| \mathbf{y}_0 \|_n^2 = 0$ 或 $\mathbf{y}_0 = \mathbf{0}$ 时,因而 $\mathbf{y} = \hat{\mathbf{y}}$ (见练习 3.1)。

这就完成了引理的证明。

3.2 新息序列

为了应用数据信息,需要一个“正交化”过程。

定义 3.1 给定一个 q 维随机向量数据序列 $\{\mathbf{v}_j\}, j=0,1,\dots,k$ 。 $\{\mathbf{v}_j\}$ 的新息序列 $\{\mathbf{z}_j\}, j=0,1,\dots,k$ (即,通过改变原始数据序列 $\{\mathbf{v}_j\}$ 得到的序列)定义为

$$\mathbf{z}_j = \mathbf{v}_j - \mathbf{C}_j \hat{\mathbf{y}}_{j-1}, \quad j = 0, 1, \dots, k \quad (3.4)$$

式中 $\hat{\mathbf{y}}_{-1} = \mathbf{0}$,且

$$\hat{\mathbf{y}}_{j-1} = \sum_{i=0}^{j-1} \hat{\mathbf{P}}_{j-1,i} \mathbf{v}_i \in Y(\mathbf{v}_0, \dots, \mathbf{v}_{j-1}), \quad j = 1, \dots, k$$

其中 $q \times n$ 阶矩阵 \mathbf{C}_j 是式(3.1)的观测矩阵。选择 $n \times q$ 阶矩阵 $\hat{\mathbf{P}}_{j-1,i}$,使得 $\hat{\mathbf{y}}_{j-1}$ 是最小化问题(3.2)中用 $Y(\mathbf{v}_0, \dots, \mathbf{v}_{j-1})$ 替代 $Y(\mathbf{w}_0, \dots, \mathbf{w}_r)$ 时的解。 ■

首先给出新息序列的相关特性。

引理 3.2 $\{\mathbf{v}_j\}$ 的新息序列 $\{\mathbf{z}_j\}$ 满足以下特性:

$$\langle \mathbf{z}_j, \mathbf{z}_l \rangle = (\mathbf{R}_l + \mathbf{C}_l \| \mathbf{x}_l - \hat{\mathbf{y}}_{l-1} \|_n^2 \mathbf{C}_l^T) \delta_{jl}$$

式中, $\mathbf{R}_l = \text{Var}(\boldsymbol{\eta}_l) > \mathbf{0}$ 。 ■

为了方便,设

$$\hat{e}_j = \mathbf{C}_j(x_j - \hat{y}_{j-1}) \quad (3.5)$$

为证明该引理,首先注意到:

$$z_j = \hat{e}_j + \underline{\eta}_j \quad (3.6)$$

式中 $\{\underline{\eta}_j\}$ ^①是观测噪声序列,且对所有的 $l \geq j$,有

$$\langle \underline{\eta}_l, \hat{e}_j \rangle = \mathbf{O}_{q \times q} \quad (3.7)$$

显然,式(3.6)可由式(3.4)、式(3.5)和观测方程(3.1)得到。式(3.7)的证明留给读者作为练习(见练习3.2)。

当 $j=l$ 时,顺次应用式(3.6)、式(3.7)和式(3.5),可得

$$\begin{aligned} \langle z_l, z_l \rangle &= \langle \hat{e}_l + \underline{\eta}_l, \hat{e}_l + \underline{\eta}_l \rangle \\ &= \langle \hat{e}_l, \hat{e}_l \rangle + \langle \underline{\eta}_l, \underline{\eta}_l \rangle \\ &= \mathbf{C}_l \|x_l - \hat{y}_{l-1}\|_n^2 \mathbf{C}_l^\top + \mathbf{R}_l \end{aligned}$$

对于 $j \neq l$,因为 $\langle \hat{e}_l, \hat{e}_j \rangle^\top = \langle \hat{e}_j, \hat{e}_l \rangle$,不失一般性可以假设 $j > l$ 。根据式(3.6)、式(3.7)和引理3.1,有

$$\begin{aligned} \langle z_j, z_l \rangle &= \langle \hat{e}_j, \hat{e}_l \rangle + \langle \hat{e}_j, \underline{\eta}_l \rangle + \langle \underline{\eta}_j, \hat{e}_l \rangle + \langle \underline{\eta}_j, \underline{\eta}_l \rangle \\ &= \langle \hat{e}_j, \hat{e}_l + \underline{\eta}_l \rangle \\ &= \langle \hat{e}_j, z_l \rangle \\ &= \langle \hat{e}_j, v_l - \mathbf{C}_l \hat{y}_{l-1} \rangle \\ &= \left\langle \mathbf{C}_j(x_j - \hat{y}_{j-1}), v_l - \mathbf{C}_l \sum_{i=0}^{l-1} \hat{\mathbf{P}}_{l-1,i} v_i \right\rangle \\ &= \mathbf{C}_j \langle x_j - \hat{y}_{j-1}, v_l \rangle - \mathbf{C}_j \sum_{i=0}^{l-1} \langle x_j - \hat{y}_{j-1}, v_i \rangle \hat{\mathbf{P}}_{l-1,i}^\top \mathbf{C}_l^\top \\ &= \mathbf{O}_{q \times q} \end{aligned}$$

这就完成了引理的证明。

因为 $\mathbf{R}_j > \mathbf{0}$,引理3.2说明 $\{z_j\}$ 是非零正交向量序列,可通过下式归一化:

$$e_j = \|z_j\|_q^{-1} z_j \quad (3.8)$$

这样,对于所有的 i 和 j ,在满足 $\langle e_i, e_j \rangle = \delta_{ij} \mathbf{I}_q$ 的意义下, $\{e_j\}$ 是正交序列。更进一步,很明显有

$$Y(e_0, \dots, e_k) = Y(v_0, \dots, v_k) \quad (3.9)$$

(见练习3.3)。

① 原文为 $\{\eta_k\}$,疑有误。

3.3 最小方差估计

下面基于相对于“正交”序列 $\{e_j\}$ 的“傅里叶级数展开”，来介绍状态向量 x_k 的最小方差估计 \check{x}_k ：

$$\check{x}_k = \sum_{i=0}^k \langle x_k, e_i \rangle e_i \quad (3.10)$$

因为

$$\langle \check{x}_k, e_j \rangle = \sum_{i=0}^k \langle x_k, e_i \rangle \langle e_i, e_j \rangle = \langle x_k, e_j \rangle$$

有

$$\langle x_k - \check{x}_k, e_j \rangle = \mathbf{O}_{n \times q}, \quad j = 0, 1, \dots, k \quad (3.11)$$

根据练习 3.3，有

$$\langle x_k - \check{x}_k, v_j \rangle = \mathbf{O}_{n \times q}, \quad j = 0, 1, \dots, k \quad (3.12)$$

从而根据引理 3.1，有

$$\text{tr} \| x_k - \check{x}_k \|_n^2 = \min \{ \text{tr} \| x_k - y \|_n^2 : y \in Y(v_0, \dots, v_k) \}$$

即， \check{x}_k 是 x_k 的最小方差估计。

3.4 卡尔曼滤波方程

本节主要推导卡尔曼滤波方程。从假设 2.1，可得

$$\langle \underline{\xi}_{k-1}, e_j \rangle = \mathbf{O}_{n \times q}, \quad j = 0, 1, \dots, k-1$$

(见练习 3.4)，因此

$$\begin{aligned} \check{x}_k &= \sum_{j=0}^k \langle x_k, e_j \rangle e_j \\ &= \sum_{j=0}^{k-1} \langle x_k, e_j \rangle e_j + \langle x_k, e_k \rangle e_k \\ &= \sum_{j=0}^{k-1} \{ \langle A_{k-1} x_{k-1}, e_j \rangle e_j + \langle \Gamma_{k-1} \underline{\xi}_{k-1}, e_j \rangle e_j \} + \langle x_k, e_k \rangle e_k \\ &= A_{k-1} \sum_{j=0}^{k-1} \langle x_{k-1}, e_j \rangle e_j + \langle x_k, e_k \rangle e_k \\ &= A_{k-1} \check{x}_{k-1} + \langle x_k, e_k \rangle e_k \end{aligned}$$

引进定义

$$\check{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \check{\mathbf{x}}_{k-1} \quad (3.13)$$

式中, $\check{\mathbf{x}}_{k-1} := \check{\mathbf{x}}_{k-1|k-1}$, 则

$$\check{\mathbf{x}}_k = \check{\mathbf{x}}_{k|k} = \check{\mathbf{x}}_{k|k-1} + \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k \quad (3.14)$$

若存在一个 $n \times q$ 阶常值矩阵 $\check{\mathbf{G}}_k$, 使得

$$\langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k = \check{\mathbf{G}}_k (\mathbf{v}_k - \mathbf{C}_k \check{\mathbf{x}}_{k|k-1})$$

即可得到卡尔曼滤波的“预测-校正”公式。为了完成这个任务, 考虑随机向量 $(\mathbf{v}_k - \mathbf{C}_k \check{\mathbf{x}}_{k|k-1})$, 并得到下面的引理。

引理 3.3 对于 $j=0, 1, \dots, k$,

$$\langle \mathbf{v}_k - \mathbf{C}_k \check{\mathbf{x}}_{k|k-1}, \mathbf{e}_j \rangle = \| \mathbf{z}_k \|_q \delta_{kj} \quad \blacksquare$$

证明如下。

首先有

$$\langle \hat{\mathbf{y}}_j, \mathbf{z}_k \rangle = \mathbf{O}_{n \times q}, \quad j = 0, 1, \dots, k-1 \quad (3.15)$$

(见练习 3.4)。

根据式(3.14)、式(3.11)和式(3.15), 可得

$$\begin{aligned} \langle \mathbf{v}_k - \mathbf{C}_k \check{\mathbf{x}}_{k|k-1}, \mathbf{e}_k \rangle &= \langle \mathbf{v}_k - \mathbf{C}_k (\check{\mathbf{x}}_{k|k} - \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k), \mathbf{e}_k \rangle \\ &= \langle \mathbf{v}_k, \mathbf{e}_k \rangle - \mathbf{C}_k \{ \langle \check{\mathbf{x}}_{k|k}, \mathbf{e}_k \rangle - \langle \mathbf{x}_k, \mathbf{e}_k \rangle \} \\ &= \langle \mathbf{v}_k, \mathbf{e}_k \rangle - \mathbf{C}_k \langle \check{\mathbf{x}}_{k|k} - \mathbf{x}_k, \mathbf{e}_k \rangle \\ &= \langle \mathbf{v}_k, \mathbf{e}_k \rangle \\ &= \langle \mathbf{z}_k + \mathbf{C}_k \hat{\mathbf{y}}_{k-1}, \| \mathbf{z}_k \|_q^{-1} \mathbf{z}_k \rangle \\ &= \langle \mathbf{z}_k, \mathbf{z}_k \rangle \| \mathbf{z}_k \|_q^{-1} + \mathbf{C}_k \langle \hat{\mathbf{y}}_{k-1}, \mathbf{z}_k \rangle \| \mathbf{z}_k \|_q^{-1} \\ &= \| \mathbf{z}_k \|_q \end{aligned}$$

根据式(3.14)、式(3.11)和式(3.7), 对于 $j=0, 1, \dots, k-1$, 有

$$\begin{aligned} \langle \mathbf{v}_k - \mathbf{C}_k \check{\mathbf{x}}_{k|k-1}, \mathbf{e}_j \rangle &= \langle \mathbf{C}_k \mathbf{x}_k + \underline{\boldsymbol{\eta}}_k - \mathbf{C}_k (\check{\mathbf{x}}_{k|k} - \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k), \mathbf{e}_j \rangle \\ &= \mathbf{C}_k \langle \mathbf{x}_k - \check{\mathbf{x}}_{k|k}, \mathbf{e}_j \rangle + \langle \underline{\boldsymbol{\eta}}_k, \mathbf{e}_j \rangle + \mathbf{C}_k \langle \mathbf{x}_k, \mathbf{e}_k \rangle \langle \mathbf{e}_k, \mathbf{e}_j \rangle \\ &= \mathbf{O}_{q \times q} \end{aligned}$$

这就完成了引理的证明。

根据练习 3.3 和定义 $\check{x}_{k-1} = \check{x}_{k-1|k-1}$, q 维随机向量 $(v_k - C_k \check{x}_{k|k-1})$ 可以表示为 $\sum_{i=0}^k M_i e_i$, 式中 M_i 为 $q \times q$ 阶常值矩阵。

现在对于 $j = 0, 1, \dots, k$, 根据引理 3.3, 有

$$\left\langle \sum_{i=0}^k M_i e_i, e_j \right\rangle = \| z_k \|_q \delta_{kj}$$

从而可得 $M_0 = M_1 = \dots = M_{k-1} = \mathbf{0}$, 且 $M_k = \| z_k \|_q e_k$ 。因此,

$$v_k - C_k \check{x}_{k|k-1} = M_k e_k = \| z_k \|_q e_k$$

定义

$$\check{G}_k = \langle x_k, e_k \rangle \| z_k \|_q^{-1}$$

可得

$$\langle x_k, e_k \rangle e_k = \check{G}_k (v_k - C_k \check{x}_{k|k-1})$$

结合式(3.14), 给出“预测-校正”方程:

$$\check{x}_{k|k} = \check{x}_{k|k-1} + \check{G}_k (v_k - C_k \check{x}_{k|k-1}) \quad (3.16)$$

我们再次强调通过选择合适的初始估计值, $\check{x}_{k|k}$ 是 x_k 的无偏估计。事实上,

$$x_k - \check{x}_{k|k} = A_{k-1} x_{k-1} + \Gamma_{k-1} \xi_{k-1} - A_{k-1} \check{x}_{k-1|k-1} - \check{G}_k (v_k - C_k A_{k-1} \check{x}_{k-1|k-1})$$

将 $v_k = C_k x_k + \eta_k = C_k A_{k-1} x_{k-1} + C_k \Gamma_{k-1} \xi_{k-1} + \eta_k$ 代入上式, 有

$$x_k - \check{x}_{k|k} = (I - \check{G}_k C_k) A_{k-1} (x_{k-1} - \check{x}_{k-1|k-1}) + (I - \check{G}_k C_k) \Gamma_{k-1} \xi_{k-1} - \check{G}_k \eta_k \quad (3.17)$$

因为噪声序列是零均值的, 则

$$E(x_k - \check{x}_{k|k}) = (I - \check{G}_k C_k) A_{k-1} E(x_{k-1} - \check{x}_{k-1|k-1})$$

所以

$$E(x_k - \check{x}_{k|k}) = (I - \check{G}_k C_k) A_{k-1} \cdots (I - \check{G}_1 C_1) A_0 E(x_0 - \check{x}_{0|0})$$

如果我们假设

$$\check{x}_{0|0} = E(x_0) \quad (3.18)$$

则对于所有的 k , $E(\mathbf{x}_k - \check{\mathbf{x}}_{k|k}) = \mathbf{0}$ 也就是 $E(\check{\mathbf{x}}_{k|k}) = E(\mathbf{x}_k)$, 即 $\check{\mathbf{x}}_{k|k}$ 是 \mathbf{x}_k 的无偏估计。

现在剩下的工作是推导 $\check{\mathbf{G}}_k$ 的递推公式。根据式(3.12)和式(3.17), 可推导出

$$\begin{aligned} \mathbf{0} &= \langle \mathbf{x}_k - \check{\mathbf{x}}_{k|k}, \mathbf{v}_k \rangle \\ &= \langle (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \mathbf{A}_{k-1} (\mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}) + (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \underline{\boldsymbol{\Gamma}}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} - \check{\mathbf{G}}_k \underline{\boldsymbol{\eta}}_k, \\ &\quad \mathbf{C}_k \mathbf{A}_{k-1} ((\mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}) + \check{\mathbf{x}}_{k-1|k-1}) + \mathbf{C}_k \underline{\boldsymbol{\Gamma}}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} + \underline{\boldsymbol{\eta}}_k \rangle \quad (3.19) \\ &= (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \mathbf{A}_{k-1} \| \mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1} \|_n^2 \mathbf{A}_{k-1}^\top \mathbf{C}_k^\top \\ &\quad + (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \underline{\boldsymbol{\Gamma}}_{k-1} \underline{\boldsymbol{Q}}_{k-1} \underline{\boldsymbol{\Gamma}}_{k-1}^\top \mathbf{C}_k^\top - \check{\mathbf{G}}_k \mathbf{R}_k \end{aligned}$$

其中使用了引理 3.1 的结论, $\langle \mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}, \check{\mathbf{x}}_{k-1|k-1} \rangle = \mathbf{O}_{n \times n}$, 且由于对于 $j = 0, \dots, k$, 有

$$\begin{aligned} \langle \mathbf{x}_k, \underline{\boldsymbol{\xi}}_k \rangle &= \mathbf{O}_{n \times n} & \langle \check{\mathbf{x}}_{k|k}, \underline{\boldsymbol{\xi}}_j \rangle &= \mathbf{O}_{n \times n} \\ \langle \mathbf{x}_k, \underline{\boldsymbol{\eta}}_j \rangle &= \mathbf{O}_{n \times q} & \langle \check{\mathbf{x}}_{k-1|k-1}, \underline{\boldsymbol{\eta}}_k \rangle &= \mathbf{O}_{n \times q} \end{aligned} \quad (3.20)$$

(见练习 3.5)。定义

$$\mathbf{P}_{k,k} = \| \mathbf{x}_k - \check{\mathbf{x}}_{k|k} \|_n^2$$

和

$$\mathbf{P}_{k,k-1} = \| \mathbf{x}_k - \check{\mathbf{x}}_{k|k-1} \|_n^2$$

由练习 3.5, 有

$$\mathbf{P}_{k,k-1} = \| \mathbf{A}_{k-1} \mathbf{x}_{k-1} + \underline{\boldsymbol{\Gamma}}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} - \mathbf{A}_{k-1} \check{\mathbf{x}}_{k-1|k-1} \|_n^2$$

$$= \mathbf{A}_{k-1} \| \mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1} \|_n^2 \mathbf{A}_{k-1}^\top + \underline{\boldsymbol{\Gamma}}_{k-1} \underline{\boldsymbol{Q}}_{k-1} \underline{\boldsymbol{\Gamma}}_{k-1}^\top$$

即

$$\mathbf{P}_{k,k-1} = \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^\top + \underline{\boldsymbol{\Gamma}}_{k-1} \underline{\boldsymbol{Q}}_{k-1} \underline{\boldsymbol{\Gamma}}_{k-1}^\top \quad (3.21)$$

另一方面, 根据式(3.19), 可得

$$(\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^\top \mathbf{C}_k^\top + (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \underline{\boldsymbol{\Gamma}}_{k-1} \underline{\boldsymbol{Q}}_{k-1} \underline{\boldsymbol{\Gamma}}_{k-1}^\top \mathbf{C}_k^\top - \check{\mathbf{G}}_k \mathbf{R}_k = \mathbf{0}$$

为了从该式中得出 $\check{\mathbf{G}}_k$, 把它改写成

$$\begin{aligned} & \check{\mathbf{G}}_k [\mathbf{R}_k + \mathbf{C}_k (\mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T) \mathbf{C}_k^T] \\ &= [\mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T] \mathbf{C}_k^T \\ &= \mathbf{P}_{k,k-1} \mathbf{C}_k^T \end{aligned}$$

可得

$$\check{\mathbf{G}}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{R}_k + \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T)^{-1} \quad (3.22)$$

式中, \mathbf{R}_k 是正定的, 且 $\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T$ 是非负定的, 所以他们的和是正定的(见练习 2.2)。

我们希望用 $\mathbf{P}_{k,k-1}$ 表示 $\mathbf{P}_{k,k}$, 以便结合式(3.21), 得到递推形式。这可以通过下面的方式得到:

$$\begin{aligned} \mathbf{P}_{k,k} &= \| \mathbf{x}_k - \check{\mathbf{x}}_{k|k} \|_n^2 \\ &= \| \mathbf{x}_k - (\check{\mathbf{x}}_{k|k-1} + \check{\mathbf{G}}_k (\mathbf{v}_k - \mathbf{C}_k \check{\mathbf{x}}_{k|k-1})) \|_n^2 \\ &= \| \mathbf{x}_k - \check{\mathbf{x}}_{k|k-1} - \check{\mathbf{G}}_k (\mathbf{C}_k \mathbf{x}_k + \underline{\eta}_k) + \check{\mathbf{G}}_k \mathbf{C}_k \check{\mathbf{x}}_{k|k-1} \|_n^2 \\ &= \| (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) (\mathbf{x}_k - \check{\mathbf{x}}_{k|k-1}) - \check{\mathbf{G}}_k \underline{\eta}_k \|_n^2 \\ &= (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \| \mathbf{x}_k - \check{\mathbf{x}}_{k|k-1} \|_n^2 (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k)^T + \check{\mathbf{G}}_k \mathbf{R}_k \check{\mathbf{G}}_k^T \\ &= (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k)^T + \check{\mathbf{G}}_k \mathbf{R}_k \check{\mathbf{G}}_k^T \end{aligned}$$

其中使用了根据练习 3.5 得出的结论 $\langle \mathbf{x}_k - \check{\mathbf{x}}_{k|k-1}, \underline{\eta}_k \rangle = \mathbf{O}_{n \times q}$ 。该关系式可以应用式(3.22)做进一步简化。因为

$$\begin{aligned} (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} (\check{\mathbf{G}}_k \mathbf{C}_k)^T &= \mathbf{P}_{k,k-1} \mathbf{C}_k^T \check{\mathbf{G}}_k^T - \check{\mathbf{G}}_k \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T \check{\mathbf{G}}_k^T \\ &= \check{\mathbf{G}}_k \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T \check{\mathbf{G}}_k^T + \check{\mathbf{G}}_k \mathbf{R}_k \check{\mathbf{G}}_k^T - \check{\mathbf{G}}_k \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T \check{\mathbf{G}}_k^T \\ &= \check{\mathbf{G}}_k \mathbf{R}_k \check{\mathbf{G}}_k^T \end{aligned}$$

有

$$\begin{aligned} \mathbf{P}_{k,k} &= (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k)^T + (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} (\check{\mathbf{G}}_k \mathbf{C}_k)^T \\ &= (\mathbf{I} - \check{\mathbf{G}}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \end{aligned} \quad (3.23)$$

综合式(3.13)、式(3.16)、式(3.18)、式(3.21)、式(3.22)和式(3.23), 及

$$\mathbf{P}_{0,0} = \| \mathbf{x}_0 - \check{\mathbf{x}}_{0|0} \|_n^2 = \text{Var}(\mathbf{x}_0) \quad (3.24)$$

我们就得到了与第2章的推导相一致的卡尔曼滤波方程。即 $\dot{\hat{x}}_{k|k} = \hat{x}_{k|k}$ 、 $\dot{\hat{x}}_{k|k-1} = \hat{x}_{k|k-1}$ 和 $\dot{G}_k = G_k$, 归纳如下:

$$\left\{ \begin{array}{l} \mathbf{P}_{0,0} = \text{Var}(\mathbf{x}_0) \\ \mathbf{P}_{k,k-1} = \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T \\ \mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ \mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \\ \hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \\ \hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \\ k = 1, 2, \dots \end{array} \right. \quad (3.25)$$

当然, 针对常规线性确定性/随机系统:

$$\left\{ \begin{array}{l} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \boldsymbol{\Gamma}_k \underline{\xi}_k \\ \mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k + \eta_k \end{array} \right.$$

在没有假定矩阵 \mathbf{A}_k 、 $\text{Var}(\underline{\xi}_{k,j})$ 等可逆的条件下, 依然可以得到2.4节推导的卡尔曼滤波方程式(2.18)(见练习3.6)。

3.5 实时跟踪

为了介绍式(3.25)描述的卡尔曼滤波算法的应用, 考虑一个实时跟踪的例子。设 $\mathbf{x}(t), 0 \leq t \leq \infty$, 表示一个飞行器在三维空间的轨迹, 其中 t 表示时间变量(见图3.1)。以采样周期 $h > 0$ 采样, 并量化得到该向量值函数的离散化形式:

$$\mathbf{x}_k \doteq \mathbf{x}(kh), \quad k = 0, 1, \dots$$

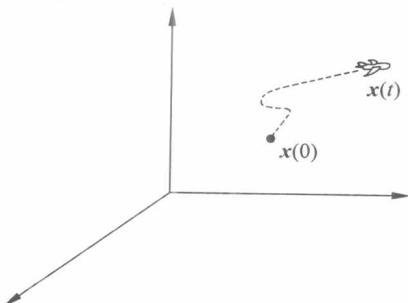


图3.1 飞行器三维空间轨迹

从实用的角度,假设 $\mathbf{x}(t)$ 有连续的一阶和二阶导数,分别表示为 $\dot{\mathbf{x}}(t)$ 和 $\ddot{\mathbf{x}}(t)$,则对于小的 h 值,位置和速度向量 \mathbf{x}_k 和 $\dot{\mathbf{x}}_k \doteq \dot{\mathbf{x}}(kh)$ 满足下面的方程:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{x}_k + h \dot{\mathbf{x}}_k + \frac{1}{2} h^2 \ddot{\mathbf{x}}_k \\ \dot{\mathbf{x}}_{k+1} = \dot{\mathbf{x}}_k + h \ddot{\mathbf{x}}_k, \end{cases}$$

式中 $\ddot{\mathbf{x}}_k \doteq \ddot{\mathbf{x}}(kh)$ 且 $k=0, 1, \dots$ 。

而在很多应用中,每个时刻的飞行器位置(向量)是可以观测的,观测值为 $\mathbf{v}_k = \mathbf{C}\mathbf{x}_k$, $\mathbf{C} = [1 \ 0 \ 0]$ 。为了简化讨论,只考虑下面的跟踪模型(见练习 3.8):

$$\begin{cases} \begin{bmatrix} \mathbf{x}_{k+1}[1] \\ \mathbf{x}_{k+1}[2] \\ \mathbf{x}_{k+1}[3] \end{bmatrix} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_k[1] \\ \mathbf{x}_k[2] \\ \mathbf{x}_k[3] \end{bmatrix} + \begin{bmatrix} \xi_k[1] \\ \xi_k[2] \\ \xi_k[3] \end{bmatrix} \\ v_k = [1 \ 0 \ 0] \begin{bmatrix} \mathbf{x}_k[1] \\ \mathbf{x}_k[2] \\ \mathbf{x}_k[3] \end{bmatrix} + \eta_k \end{cases} \quad (3.26)$$

这里, $\underline{\xi}_k := [\xi_k[1] \ \xi_k[2] \ \xi_k[3]]^T$, $\{\underline{\xi}_k\}$ 和 $\{\eta_k\}$ 为满足下式的零均值高斯白噪声序列:

$$\begin{aligned} E(\underline{\xi}_k) &= \mathbf{0} & E(\eta_k) &= 0 \\ E(\underline{\xi}_k \underline{\xi}_l^T) &= Q_k \delta_{kl} & E(\eta_k \eta_l) &= r_k \delta_{kl} & E(\underline{\xi}_k \eta_l) &= \mathbf{0} \\ E(x_0 \underline{\xi}_k^T) &= \mathbf{0} & E(x_0 \eta_k) &= \mathbf{0} \end{aligned}$$

式中 Q_k 是非负定对称矩阵,且对所有的 k ,有 $r_k > 0$ 。进一步假设初始条件 $E(x_0)$ 和 $\text{Var}(x_0)$ 为已知。对于这个跟踪模型,卡尔曼滤波算法可以表示为:设 $\mathbf{P}_k := \mathbf{P}_{k,k}$,且 $\mathbf{P}[i,j]$ 表示 \mathbf{P} 的第 (i,j) 个元素,则有

$$\begin{aligned} \mathbf{P}_{k,k-1}[1,1] &= \mathbf{P}_{k-1}[1,1] + 2h\mathbf{P}_{k-1}[1,2] + h^2\mathbf{P}_{k-1}[1,3] + h^2\mathbf{P}_{k-1}[2,2] \\ &\quad + h^3\mathbf{P}_{k-1}[2,3] + \frac{h^4}{4}\mathbf{P}_{k-1}[3,3] + Q_{k-1}[1,1] \end{aligned}$$

$$\begin{aligned} \mathbf{P}_{k,k-1}[1,2] &= \mathbf{P}_{k-1}[2,1] \\ &= \mathbf{P}_{k-1}[1,2] + h\mathbf{P}_{k-1}[1,3] + h\mathbf{P}_{k-1}[2,2] + \frac{3h^2}{2}\mathbf{P}_{k-1}[2,3] \\ &\quad + \frac{h^3}{2}\mathbf{P}_{k-1}[3,3] + Q_{k-1}[1,2] \end{aligned}$$

$$\mathbf{P}_{k,k-1}[2,2] = \mathbf{P}_{k-1}[2,2] + 2h\mathbf{P}_{k-1}[2,3] + h^2\mathbf{P}_{k-1}[3,3] + Q_{k-1}[2,2]$$

$$\mathbf{P}_{k,k-1}[1,3] = \mathbf{P}_{k-1}[3,1]$$

$$= \mathbf{P}_{k-1}[1,3] + h\mathbf{P}_{k-1}[2,3] + \frac{h^2}{2}\mathbf{P}_{k-1}[3,3] + \mathbf{Q}_{k-1}[1,3]$$

$$\mathbf{P}_{k,k-1}[2,3] = \mathbf{P}_{k,k-1}[3,2]$$

$$= \mathbf{P}_{k-1}[2,3] + h\mathbf{P}_{k-1}[3,3] + \mathbf{Q}_{k-1}[2,3]$$

$$\mathbf{P}_{k,k-1}[3,3] = \mathbf{P}_{k-1}[3,3] + \mathbf{Q}_{k-1}[3,3]$$

且 $\mathbf{P}_{0,0} = \text{Var}(\mathbf{x}_0)$

$$\mathbf{G}_k = \frac{1}{\mathbf{P}_{k,k-1}[1,1] + r_k} \begin{bmatrix} \mathbf{P}_{k,k-1}[1,1] \\ \mathbf{P}_{k,k-1}[1,2] \\ \mathbf{P}_{k,k-1}[1,3] \end{bmatrix}$$

$$\mathbf{P}_k = \mathbf{P}_{k,k-1} - \frac{1}{\mathbf{P}_{k,k-1}[1,1] + r_k}$$

$$\cdot \begin{bmatrix} \mathbf{P}_{k,k-1}^2[1,1] & \mathbf{P}_{k,k-1}[1,1]\mathbf{P}_{k,k-1}[1,2] & \mathbf{P}_{k,k-1}[1,1]\mathbf{P}_{k,k-1}[1,3] \\ \mathbf{P}_{k,k-1}[1,1]\mathbf{P}_{k,k-1}[1,2] & \mathbf{P}_{k,k-1}^2[1,2] & \mathbf{P}_{k,k-1}[1,2]\mathbf{P}_{k,k-1}[1,3] \\ \mathbf{P}_{k,k-1}[1,1]\mathbf{P}_{k,k-1}[1,3] & \mathbf{P}_{k,k-1}[1,2]\mathbf{P}_{k,k-1}[1,3] & \mathbf{P}_{k,k-1}^2[1,3] \end{bmatrix}$$

卡尔曼滤波算法由下式给出：

$$\begin{bmatrix} \hat{\mathbf{x}}_{k|k}[1] \\ \hat{\mathbf{x}}_{k|k}[2] \\ \hat{\mathbf{x}}_{k|k}[3] \end{bmatrix} = \begin{bmatrix} 1 - \mathbf{G}_k[1](1 - \mathbf{G}_k[1])h(1 - \mathbf{G}_k[1])h^2/2 \\ -\mathbf{G}_k[2] & 1 - h\mathbf{G}_k[2] & h - h^2\mathbf{G}_k[2]/2 \\ -\mathbf{G}_k[3] & -h\mathbf{G}_k[3] & 1 - h^2\mathbf{G}_k[3]/2 \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{x}}_{k-1|k-1}[1] \\ \hat{\mathbf{x}}_{k-1|k-1}[2] \\ \hat{\mathbf{x}}_{k-1|k-1}[3] \end{bmatrix} + \begin{bmatrix} \mathbf{G}_k[1] \\ \mathbf{G}_k[2] \\ \mathbf{G}_k[3] \end{bmatrix} v_k \quad (3.27)$$

式中, $\hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0)$ 。

练习

3.1 令 $\mathbf{A} \neq 0$ 是非负定对称常值矩阵。证明 $\text{tr}\mathbf{A} > 0$ 。(提示：将 \mathbf{A} 分解为 $\mathbf{A} = \mathbf{B}\mathbf{B}^T$, 其中 $\mathbf{B} \neq 0$)。

3.2 令

$$\hat{\mathbf{e}}_j = \mathbf{C}_j(\mathbf{x}_j - \hat{\mathbf{y}}_{j-1}) = \mathbf{C}_j \left(\mathbf{x}_j - \sum_{i=0}^{j-1} \hat{\mathbf{P}}_{j-1,i} \mathbf{v}_i \right)$$

式中 $\hat{\mathbf{P}}_{j-1,i}$ 是常值矩阵。应用假设 2.1 证明：对于所有 $l \geq j$, 有

$$\langle \boldsymbol{\eta}_l, \hat{\mathbf{e}}_j \rangle = \mathbf{O}_{q \times q}$$

3.3 对随机向量 $\mathbf{w}_0, \dots, \mathbf{w}_r$, 定义: $\mathbf{Y}(\mathbf{w}_0, \dots, \mathbf{w}_r) = \left\{ \mathbf{y}: \mathbf{y} = \sum_{i=0}^r \mathbf{P}_i \mathbf{w}_i, \mathbf{P}_0, \dots, \mathbf{P}_r \text{ 为常值矩阵} \right\}$ 。

令

$$\mathbf{z}_j = \mathbf{v}_j - \mathbf{C}_j \sum_{i=0}^{j-1} \hat{\mathbf{P}}_{j-1,i} \mathbf{v}_i$$

和式(3.4)所定义的一样,且 $\mathbf{e}_j = \|\mathbf{z}_j\|^{-1} \mathbf{z}_j$ 。证明: $Y(\mathbf{e}_0, \dots, \mathbf{e}_k) = Y(\mathbf{v}_0, \dots, \mathbf{v}_k)$ 。

3.4 令

$$\hat{\mathbf{y}}_{j-1} = \sum_{i=0}^{j-1} \hat{\mathbf{P}}_{j-1,i} \mathbf{v}_i \quad \text{和} \quad \mathbf{z}_j = \mathbf{v}_j - \mathbf{C}_j \sum_{i=0}^{j-1} \hat{\mathbf{P}}_{j-1,i} \mathbf{v}_i$$

证明: $\langle \hat{\mathbf{y}}_j, \mathbf{z}_k \rangle = \mathbf{O}_{n \times q}, j=0,1,\dots,k-1$ 。

3.5 令 \mathbf{e}_j 的定义与练习 3.3 一致,另外与式(3.10)同样定义:

$$\check{\mathbf{x}}_k = \sum_{i=0}^k \langle \mathbf{x}_k, \mathbf{e}_i \rangle \mathbf{e}_i$$

对于所有的 $j=0,1,\dots,k$,证明:

$$\langle \mathbf{x}_k, \xi_k \rangle = \mathbf{O}_{n \times n} \quad \langle \check{\mathbf{x}}_{k|k}, \xi_j \rangle = \mathbf{O}_{n \times n}$$

$$\langle \mathbf{x}_k, \eta_j \rangle = \mathbf{O}_{n \times q} \quad \langle \check{\mathbf{x}}_{k-1|k-1}, \eta_k \rangle = \mathbf{O}_{n \times q}$$

3.6 对于线性确定/随机系统:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \Gamma_k \xi_k \\ \mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k + \eta_k \end{cases}$$

式中, $\{\mathbf{u}_k\}$ 为给定的 m 维确定性控制输入向量序列, $1 \leq m \leq n$, 如果满足假设 2.1, 推导该模型的卡尔曼滤波算法。

3.7 对于一个简化的雷达跟踪模型,天线发射大幅值窄带脉冲信号。该脉冲信号以光速 c 传播,由被跟踪飞行器反射。雷达天线接收反射信号,存在一个时间差 Δt 。从雷达到目标的距离可以由 $d = c \Delta t / 2$ 给出。脉冲信号以周期 h 发射。假设目标以速度 w 运动,且速度受到的随机干扰为 $\xi \sim N(0, q)$,则距离满足差分方程: $d_{k+1} = d_k + h(w_k + \xi_k)$ 。

假设应用公式 $d = c \Delta t / 2$ 获得的测量距离包含固有误差 Δd 且被噪声 η 污染,其中 $\eta \sim N(0, r)$,则: $v_k = d_k + \Delta d_k + \eta_k$ 。

假设雷达与目标的初始距离是 d_0 ,且与 ξ_k 和 η_k 独立,并且 $\{\xi_k\}$ 和 $\{\eta_k\}$ 也相互独立(见图 3.2)。推导该雷达跟踪系统距离估计的卡尔曼滤波算法。

3.8 一个雷达跟踪线性随机系统描述如下,其中雷达被安置在原点(见图 3.3)。设 Σ 、 ΔA 和 ΔE 分别为目标的距离、方位角误差和仰角误差。考虑 Σ 、 ΔA 和 ΔE 是时间的函数,并且一阶和二阶导数分别表示为 $\dot{\Sigma}$ 、 $\dot{\Delta A}$ 、 $\ddot{\Delta E}$ 、 $\dot{\Sigma}$ 、 $\dot{\Delta A}$ 和 $\ddot{\Delta E}$ 。 $h > 0$

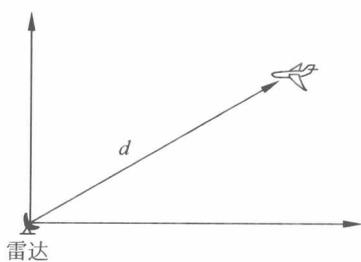


图 3.2 雷达跟踪示例

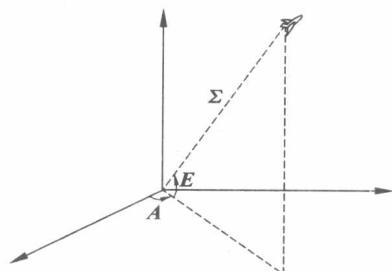


图 3.3 目标位置示意图

为采样时间单位,并记 $\Sigma_k = \Sigma(kh)$ 、 $\dot{\Sigma}_k = \dot{\Sigma}(kh)$ 、 $\ddot{\Sigma}_k = \ddot{\Sigma}(kh)$ 等。二阶泰勒多项式展开后,雷达跟踪模型可用如下的线性随机状态空间描述:

$$\begin{cases} \underline{x}_{k+1} = \tilde{A}\underline{x}_k + \Gamma_k \underline{\xi}_k \\ \underline{v}_k = \tilde{C}\underline{x}_k + \underline{\eta}_k \end{cases}$$

式中

$$\underline{x}_k = [\Sigma_k \quad \dot{\Sigma}_k \quad \ddot{\Sigma}_k \quad \Delta A_k \quad \Delta \dot{A}_k \quad \Delta \ddot{A}_k \quad \Delta E_k \quad \Delta \dot{E}_k \quad \Delta \ddot{E}_k]^T$$

$$\tilde{A} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \\ & & 1 & h & h^2/2 \\ & & 0 & 1 & h \\ & & 0 & 0 & 1 \\ & & & 1 & h & h^2/2 \\ & & & 0 & 1 & h \\ & & & 0 & 0 & 1 \end{bmatrix}$$

$$\tilde{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

其中 $\{\underline{\xi}_k\}$ 和 $\{\underline{\eta}_k\}$ 是相互独立的零均值高斯白噪声序列,且 $\text{Var}(\underline{\xi}_k) = Q_k$ 和 $\text{Var}(\underline{\eta}_k) = R_k$ 。记

$$\boldsymbol{\Gamma}_k = \begin{bmatrix} \boldsymbol{\Gamma}_k^1 \\ & \boldsymbol{\Gamma}_k^2 \\ & & \boldsymbol{\Gamma}_k^3 \end{bmatrix}$$

$$\boldsymbol{Q}_k = \begin{bmatrix} \boldsymbol{Q}_k^1 \\ & \boldsymbol{Q}_k^2 \\ & & \boldsymbol{Q}_k^3 \end{bmatrix}$$

$$\boldsymbol{R}_k = \begin{bmatrix} \boldsymbol{R}_k^1 \\ & \boldsymbol{R}_k^2 \\ & & \boldsymbol{R}_k^3 \end{bmatrix}$$

其中对于 $i=1,2,3$, $\boldsymbol{\Gamma}_k^i$ 是 3×3 阶子矩阵, \boldsymbol{Q}_k^i 是 3×3 阶非负定对称子矩阵, \boldsymbol{R}_k^i 是 3×3 阶正定对称子矩阵。证明该系统可以分解为三个具有相似状态空间描述的子系统。

第 4 章

系统噪声和量测噪声相关的卡尔曼滤波

在第 2、3 章中, 我们研究了模型中的系统噪声和量测噪声过程不相关的卡尔曼滤波; 即, 对于 $k, l = 0, 1, \dots$, 始终假设:

$$E(\underline{\xi}_k \underline{\eta}_l^T) = \mathbf{0}$$

但在许多实际应用中, 这个假设不成立。例如飞行器惯性导航系统中, 飞行器的振动成为动态系统及机载雷达测量的共同噪声源, 系统和量测噪声序列 $\{\underline{\xi}_k\}$ 和 $\{\underline{\eta}_k\}$ 在统计学意义上是相关的, 有

$$E(\underline{\xi}_k \underline{\eta}_l^T) = \mathbf{S}_k \delta_{kl}, \quad k, l = 0, 1, \dots$$

式中, 每个 \mathbf{S}_k 都是已知的非负定矩阵。

本章主要研究上述模型的卡尔曼滤波。

4.1 仿射模型

考虑线性随机状态空间模型

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \boldsymbol{\Gamma}_k \underline{\xi}_k \\ \mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \underline{\eta}_k \end{cases}$$

式中, \mathbf{A}_k 、 \mathbf{C}_k 和 $\boldsymbol{\Gamma}_k$ 都是已知的常值矩阵, 初始状态为 \mathbf{x}_0 。

首先研究 1.3 节讨论的最小二乘估计。考虑到最小二乘估计是数据向量的线性函数; 即, 如果 $\hat{\mathbf{x}}$ 是状态向量 \mathbf{x} 基于数据 \mathbf{v} 的最小二乘估计, 则存在矩阵 \mathbf{H} ,

满足 $\hat{x} = \mathbf{H}\mathbf{v}$ 。为了研究卡尔曼滤波在系统噪声与测量噪声过程相关系统中的应用,必须把它扩展成为一个更一般的模型来确定估计量 \hat{x} 。为此,考虑下面的仿射模型:

$$\hat{x} = \mathbf{h} + \mathbf{H}\mathbf{v} \quad (4.1)$$

式中 \mathbf{h} 为 n 维常向量, \mathbf{H} 是 $n \times q$ 阶常值矩阵。式(4.1)引入了一个额外的参数向量 \mathbf{h} 。当然,状态 \mathbf{x} 的最优估计 \hat{x} 必须满足: \hat{x} 是 \mathbf{x} 的一个无偏估计,即

$$E(\hat{x}) = E(\mathbf{x}) \quad (4.2)$$

且此估计具有最小(误差)方差。

由式(4.1)可得

$$\mathbf{h} = E(\mathbf{h}) = E(\hat{x} - \mathbf{H}\mathbf{v}) = E(\hat{x}) - \mathbf{H}(E(\mathbf{v}))$$

为了满足式(4.2),必须有

$$\mathbf{h} = E(\mathbf{x}) - \mathbf{H}E(\mathbf{v}) \quad (4.3)$$

这等同于

$$\hat{x} = E(\mathbf{x}) - \mathbf{H}(E(\mathbf{v}) - \mathbf{v}) \quad (4.4)$$

另一方面,为满足最小方差要求,考虑下面的关系:

$$F(\mathbf{H}) = \text{Var}(\mathbf{x} - \hat{x}) = \|\mathbf{x} - \hat{x}\|_n^2$$

由式(4.4)和 $\|\mathbf{v}\|_q^2 = \text{Var}(\mathbf{v})$ 正定,可得

$$\begin{aligned} F(\mathbf{H}) &= \langle \mathbf{x} - \hat{x}, \mathbf{x} - \hat{x} \rangle \\ &= \langle (\mathbf{x} - E(\mathbf{x})) - \mathbf{H}(\mathbf{v} - E(\mathbf{v})), (\mathbf{x} - E(\mathbf{x})) - \mathbf{H}(\mathbf{v} - E(\mathbf{v})) \rangle \\ &= \|\mathbf{x}\|_n^2 - \mathbf{H}\langle \mathbf{v}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{v} \rangle \mathbf{H}^T + \mathbf{H} \|\mathbf{v}\|_q^2 \mathbf{H}^T \\ &= \{ \|\mathbf{x}\|_n^2 - \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1} \langle \mathbf{v}, \mathbf{x} \rangle \} \\ &\quad + \{ \mathbf{H} \|\mathbf{v}\|_q^2 \mathbf{H}^T - \mathbf{H}\langle \mathbf{v}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{v} \rangle \mathbf{H}^T + \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1} \langle \mathbf{v}, \mathbf{x} \rangle \} \\ &= \{ \|\mathbf{x}\|_n^2 - \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1} \langle \mathbf{v}, \mathbf{x} \rangle \} \\ &\quad + [\mathbf{H} - \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1}] \|\mathbf{v}\|_q^2 [\mathbf{H} - \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1}]^T \end{aligned}$$

其中,应用了 $\langle \mathbf{x}, \mathbf{v} \rangle^T = \langle \mathbf{v}, \mathbf{x} \rangle$ 和 $\text{Var}(\mathbf{v})$ 是非奇异的事实。

最小方差估计表示存在 \mathbf{H}^* 使得 $F(\mathbf{H}) \geq F(\mathbf{H}^*)$,或者说对于所有常值矩阵 \mathbf{H} , $F(\mathbf{H}) - F(\mathbf{H}^*)$ 都是非负定的。这可以通过下式简单的设定得到

$$\mathbf{H}^* = \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1} \quad (4.5)$$

这是因为

$$F(\mathbf{H}) - F(\mathbf{H}^*) = [\mathbf{H} - \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1}] \|\mathbf{v}\|_q^2 [\mathbf{H} - \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1}]^T$$

式中所有常值矩阵 \mathbf{H} 都是非负定的。 $F(\mathbf{H}) - F(\mathbf{H}^*) = \mathbf{0}$ 当且仅当 $\mathbf{H} = \mathbf{H}^*$; 在这一意义下, \mathbf{H}^* 是唯一的。故 \hat{x} 可以唯一地表示为

$$\hat{x} = \mathbf{h} + \mathbf{H}^* \mathbf{v}$$

其中 \mathbf{H}^* 由式(4.5)给出。令 $\hat{\mathbf{x}} = L(\mathbf{x}, \mathbf{v})$ 表示由数据 \mathbf{v} 得到的 \mathbf{x} 的最优估计。由式(4.4)和式(4.5), 可得“最优估计算子”满足:

$$L(\mathbf{x}, \mathbf{v}) = E(\mathbf{x}) + \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1} (\mathbf{v} - E(\mathbf{v})) \quad (4.6)$$

4.2 最优估计算子

首先, 注意到对于任意固定数据向量 \mathbf{v} , 以及所有常数矩阵 \mathbf{A}, \mathbf{B} 和状态向量 \mathbf{x}, \mathbf{y} , $L(\cdot, \mathbf{v})$ 是一个线性算子, 即

$$L(\mathbf{Ax} + \mathbf{By}, \mathbf{v}) = \mathbf{AL}(\mathbf{x}, \mathbf{v}) + \mathbf{BL}(\mathbf{y}, \mathbf{v}) \quad (4.7)$$

(见练习 4.1)。

此外, 如果状态向量是一个常数向量 \mathbf{a} , 则

$$L(\mathbf{a}, \mathbf{v}) = \mathbf{a} \quad (4.8)$$

(见练习 4.2)。

这就意味着如果 \mathbf{x} 是一个常值向量, 则 $E(\mathbf{x}) = \mathbf{x}$, 即 $\hat{\mathbf{x}} = \mathbf{x}$, 或者说估计是准确的。

我们需要了解 $L(\mathbf{x}, \mathbf{v})$ 更多的一些属性, 为此我们首先建立下面的引理。

引理 4.1 假定 \mathbf{v} 是一个给定的数据向量, 且有 $\mathbf{y} = \mathbf{h} + \mathbf{H}\mathbf{v}$, 其中 \mathbf{h} 由条件 $E(\mathbf{y}) = E(\mathbf{x})$ 决定, 使得 \mathbf{y} 由常数矩阵 \mathbf{H} 唯一确定。如果 \mathbf{x}^* 是其中这样的一个 \mathbf{y} , 使得:

$$\operatorname{tr} \|\mathbf{x} - \mathbf{x}^*\|_n^2 = \min_H \operatorname{tr} \|\mathbf{x} - \mathbf{y}\|_n^2$$

则有 $\mathbf{x}^* = \hat{\mathbf{x}}$, 式中 $\hat{\mathbf{x}} = L(\mathbf{x}, \mathbf{v})$ 由式(4.6)给出。 ■

此引理表明从相同数据 \mathbf{v} 得到的 \mathbf{x} 的最小方差估计 $\hat{\mathbf{x}}$ 和最小迹方差估计 \mathbf{x}^* 对所有仿射模型是一致的。

下面证明此引理。根据式(4.3), 有

$$\begin{aligned} \operatorname{tr} \|\mathbf{x} - \mathbf{y}\|_n^2 &= \operatorname{tr} E((\mathbf{x} - \mathbf{y})(\mathbf{x} - \mathbf{y})^\top) \\ &= E((\mathbf{x} - \mathbf{y})^\top(\mathbf{x} - \mathbf{y})) \\ &= E(\mathbf{x} - E(\mathbf{x}))^\top \mathbf{H} (\mathbf{v} - E(\mathbf{v}))^\top (\mathbf{x} - E(\mathbf{x})) - \mathbf{H} (\mathbf{v} - E(\mathbf{v})) \end{aligned}$$

取

$$\frac{\partial}{\partial \mathbf{H}} (\operatorname{tr} \|\mathbf{x} - \mathbf{y}\|_n^2) = \mathbf{0}$$

可得^①

$$\mathbf{x}^* = E(\mathbf{x}) + \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|_q^2]^{-1} (\mathbf{v} - E(\mathbf{v})) \quad (4.9)$$

这与式(4.6)得到的 $\hat{\mathbf{x}}$ 相同(见练习 4.3)。至此完成了引理的证明。

^① 原文下式疑有误, 已改。

4.3 额外数据对最优估计的影响

由引理 3.1 可知, 对于 $\hat{y} \in Y = Y(w_0, \dots, w_r)$, 当且仅当

$$\langle x - \hat{y}, w_j \rangle = \mathbf{0}_{n \times q}, \quad j = 0, 1, \dots, r$$

有

$$\text{tr} \| x - \hat{y} \|_n^2 = \min_{y \in Y} \text{tr} \| x - y \|_n^2$$

令 $Y = Y(v - E(v))$, $\hat{x} = L(x, v) = E(x) + H^*(v - E(v))$, 其中 $H^* = \langle x, v \rangle [\| v \|_q^2]^{-1}$ 。记

$$\tilde{x} = x - E(x) \quad \text{和} \quad \tilde{v} = v - E(v)$$

可得

$$\| x - \hat{x} \|_n^2 = \| (x - E(x)) - H^*(v - E(v)) \|_n^2 = \| \tilde{x} - H^* \tilde{v} \|_n^2$$

但是选择 H^* 使得对于任意 H 都有 $F(H^*) \leq F(H)$, 即对于所有的 H , 都有 $\text{tr}F(H^*) \leq \text{tr}F(H)$, 因此, 对于所有的 $y \in Y(v - E(v)) = Y(\tilde{v})$, 有:

$$\text{tr} \| \tilde{x} - H^* \tilde{v} \|_n^2 \leq \text{tr} \| \tilde{x} - y \|_n^2$$

由引理 3.1, 可知

$$\langle \tilde{x} - H^* \tilde{v}, \tilde{v} \rangle = \mathbf{0}_{n \times q}$$

由于 $E(v)$ 是一个常值, $\langle \tilde{x} - H^* \tilde{v}, E(v) \rangle = \mathbf{0}_{n \times q}$, 所以

$$\langle \tilde{x} - H^* \tilde{v}, v \rangle = \mathbf{0}_{n \times q}$$

即

$$\langle x - \hat{x}, v \rangle = \mathbf{0}_{n \times q} \tag{4.10}$$

考虑两个随机数据向量 v^1 和 v^2 , 且令

$$\begin{cases} x^\# = x - L(x, v^1) \\ v^{2\#} = v^2 - L(v^2, v^1) \end{cases} \tag{4.11}$$

则由式(4.10)和定义的最优估计算子 L , 可得

$$\langle x^\#, v^1 \rangle = \mathbf{0} \tag{4.12}$$

同理

$$\langle v^{2\#}, v^1 \rangle = \mathbf{0} \tag{4.13}$$

下面的引理对进一步研究是必不可少的。

引理 4.2 令 x 为一个状态向量, v^1, v^2 是非零有限方差观测数据向量。令

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}^1 \\ \mathbf{v}^2 \end{bmatrix}$$

则在下式的意义下, \mathbf{x} 基于数据 \mathbf{v} 的最小方差估计 $\hat{\mathbf{x}}$ 可以近似为 \mathbf{x} 基于数据 \mathbf{v}^1 的最小方差估计 $L(\mathbf{x}, \mathbf{v}^1)$:

$$\hat{\mathbf{x}} = L(\mathbf{x}, \mathbf{v}) = L(\mathbf{x}, \mathbf{v}^1) + e(\mathbf{x}, \mathbf{v}^2) \quad (4.14)$$

误差为

$$e(\mathbf{x}, \mathbf{v}^2) := L(\mathbf{x}^\# , \mathbf{v}^{2\#}) = \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle [\|\mathbf{v}^{2\#}\|^2]^{-1} \mathbf{v}^{2\#} \quad (4.15) \blacksquare$$

首先来证明式(4.15)。由于 $L(\mathbf{x}, \mathbf{v}^1)$ 是 \mathbf{x} 的一个无偏估计(见式(4.6)),

$$E(\mathbf{x}^\#) = E(\mathbf{x} - L(\mathbf{x}, \mathbf{v}^1)) = \mathbf{0}$$

同理可得, $E(\mathbf{v}^{2\#}) = \mathbf{0}$ 。因此, 根据式(4.6), 可得

$$\begin{aligned} L(\mathbf{x}^\# , \mathbf{v}^{2\#}) &= E(\mathbf{x}^\#) + \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle [\|\mathbf{v}^{2\#}\|^2]^{-1} (\mathbf{v}^{2\#} - E(\mathbf{v}^{2\#})) \\ &= \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle [\|\mathbf{v}^{2\#}\|^2]^{-1} \mathbf{v}^{2\#} \end{aligned}$$

即得到式(4.15)。接下来要证明式(4.14), 即证明

$$\mathbf{x}^0 := L(\mathbf{x}, \mathbf{v}^1) + L(\mathbf{x}^\# , \mathbf{v}^{2\#})$$

这是根据数据 \mathbf{v} 得到的 \mathbf{x} 的一个仿射无偏最小方差估计。从而, 根据 $\hat{\mathbf{x}}$ 的唯一性, 就获得 $\mathbf{x}^0 = \hat{\mathbf{x}} = L(\mathbf{x}, \mathbf{v})$ 。首先注意到

$$\begin{aligned} \mathbf{x}^0 &:= L(\mathbf{x}, \mathbf{v}^1) + L(\mathbf{x}^\# , \mathbf{v}^{2\#}) \\ &= (\mathbf{h}^1 + \mathbf{H}_1 \mathbf{v}^1) + (\mathbf{h}_2 + \mathbf{H}_2 (\mathbf{v}^2 - L(\mathbf{v}^2, \mathbf{v}^1))) \\ &= (\mathbf{h}^1 + \mathbf{H}_1 \mathbf{v}^1) + \mathbf{h}_2 + \mathbf{H}_2 (\mathbf{v}^2 - (\mathbf{h}_3 + \mathbf{H}_3 \mathbf{v}^1)) \\ &= (\mathbf{h}_1 + \mathbf{h}_2 - \mathbf{H}_2 \mathbf{h}_3) + \mathbf{H} \begin{bmatrix} \mathbf{v}^1 \\ \mathbf{v}^2 \end{bmatrix} \\ &:= \mathbf{h} + \mathbf{H} \mathbf{v} \end{aligned}$$

式中, $\mathbf{H} = [\mathbf{H}_1 - \mathbf{H}_2 \mathbf{H}_3 \quad \mathbf{H}_2]$ 。因此 \mathbf{x}^0 为 \mathbf{v} 的一个仿射变换。

其次, 由于 $E(L(\mathbf{x}, \mathbf{v}^1)) = E(\mathbf{x})$, 且 $E(L(\mathbf{x}^\# , \mathbf{v}^{2\#})) = E(\mathbf{x}^\#) = \mathbf{0}$, 可得

$$E(\mathbf{x}^0) = E(L(\mathbf{x}, \mathbf{v}^1)) + E(L(\mathbf{x}^\# , \mathbf{v}^{2\#})) = E(\mathbf{x})$$

因此, \mathbf{x}^0 为 \mathbf{x} 的一个无偏估计。

最后, 为了证明 \mathbf{x}^0 为 \mathbf{x} 的一个最小方差估计, 运用引理 4.1 和引理 3.1, 可以确定它们满足正交性:

$$\langle \mathbf{x} - \mathbf{x}^0, \mathbf{v} \rangle = \mathbf{O}_{n \times q}$$

上式的证明如下。由式(4.15)、式(4.11)、式(4.12)和式(4.13), 可得

$$\langle \mathbf{x} - \mathbf{x}^0, \mathbf{v} \rangle = \langle \mathbf{x}^\# - \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle [\|\mathbf{v}^{2\#}\|^2]^{-1} \mathbf{v}^{2\#}, \mathbf{v} \rangle$$

$$= \left\langle \mathbf{x}^\# , \begin{bmatrix} \mathbf{v}^1 \\ \mathbf{v}^2 \end{bmatrix} \right\rangle - \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle [\|\mathbf{v}^{2\#}\|^2]^{-1} \left\langle \mathbf{v}^{2\#} , \begin{bmatrix} \mathbf{v}^1 \\ \mathbf{v}^2 \end{bmatrix} \right\rangle$$

$$= \langle \mathbf{x}^\# , \mathbf{v}^2 \rangle - \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle [\| \mathbf{v}^{2\#} \|^2]^{-1} \langle \mathbf{v}^{2\#} , \mathbf{v}^2 \rangle$$

但是由于 $\mathbf{v}^2 = \mathbf{v}^{2\#} + L(\mathbf{v}^2, \mathbf{v}^1)$, 因而有

$$\langle \mathbf{v}^{2\#} , \mathbf{v}^2 \rangle = \langle \mathbf{v}^{2\#} , \mathbf{v}^{2\#} \rangle + \langle \mathbf{v}^{2\#} , L(\mathbf{v}^2, \mathbf{v}^1) \rangle$$

由上式, 通过应用式(4.6)、式(4.13), 及 $\langle \mathbf{v}^{2\#}, E(\mathbf{v}^1) \rangle = \langle \mathbf{v}^{2\#}, E(\mathbf{v}^2) \rangle = \mathbf{0}$, 可得

$$\begin{aligned} \langle \mathbf{v}^{2\#} , L(\mathbf{v}^2, \mathbf{v}^1) \rangle &= \langle \mathbf{v}^{2\#} , E(\mathbf{v}^2) + \langle \mathbf{v}^2, \mathbf{v}^1 \rangle [\| \mathbf{v}^1 \|^2]^{-1} (\mathbf{v}^1 - E(\mathbf{v}^1)) \rangle \\ &= (\langle \mathbf{v}^{2\#} , \mathbf{v}^1 \rangle - \langle \mathbf{v}^{2\#} , E(\mathbf{v}^1) \rangle) [\| \mathbf{v}^1 \|^2]^{-1} \langle \mathbf{v}^1, \mathbf{v}^2 \rangle \\ &= \mathbf{0} \end{aligned}$$

因此

$$\langle \mathbf{v}^{2\#} , \mathbf{v}^2 \rangle = \langle \mathbf{v}^{2\#} , \mathbf{v}^{2\#} \rangle$$

同理可得

$$\langle \mathbf{x}^\# , \mathbf{v}^2 \rangle = \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle$$

最终得到正交性:

$$\begin{aligned} \langle \mathbf{x} - \mathbf{x}^0, \mathbf{v} \rangle &= \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle - \langle \mathbf{x}^\# , \mathbf{v}^{2\#} \rangle [\| \mathbf{v}^{2\#} \|^2]^{-1} \langle \mathbf{v}^{2\#} , \mathbf{v}^{2\#} \rangle \\ &= \mathbf{0}_{n \times q} \end{aligned}$$

证明完毕。

4.4 卡尔曼滤波方程推导

现在开始研究系统和量测噪声相关系统的卡尔曼滤波。依然考虑下面的线性随机系统:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \boldsymbol{\Gamma}_k \boldsymbol{\xi}_k \\ \mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \boldsymbol{\eta}_k \end{cases} \quad (4.16)$$

式中 \mathbf{A}_k 、 \mathbf{C}_k 和 $\boldsymbol{\Gamma}_k$ 是已知的常值矩阵, 初始状态为 \mathbf{x}_0 。这里采用假设 2.1, 其中的不同之处是两个噪声序列 $\{\boldsymbol{\xi}_k\}$ 和 $\{\boldsymbol{\eta}_k\}$ 可能是相关的, 即: 假设 $\{\boldsymbol{\xi}_k\}$ 和 $\{\boldsymbol{\eta}_k\}$ 是零均值高斯白噪声序列, 满足:

$$E(\boldsymbol{\xi}_k \mathbf{x}_0^\top) = \mathbf{0}_{p \times n}, \quad E(\boldsymbol{\eta}_k \mathbf{x}_0^\top) = \mathbf{0}_{q \times n}$$

$$E(\boldsymbol{\xi}_k \boldsymbol{\xi}_l^\top) = \mathbf{Q}_k \delta_{kl}, \quad E(\boldsymbol{\eta}_k \boldsymbol{\eta}_l^\top) = \mathbf{R}_k \delta_{kl}$$

$$E(\boldsymbol{\xi}_k \boldsymbol{\eta}_l^\top) = \mathbf{S}_k \delta_{kl}$$

式中, \mathbf{Q}_k 、 \mathbf{R}_k 分别是已知的非负定矩阵和正定矩阵; \mathbf{S}_k 是已知的但不一定是零的非负定矩阵。

问题是应用初始信息 $E(\mathbf{x}_0)$ 和 $\text{Var}(\mathbf{x}_0)$ 并基于数据向量 $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_k$, 来确

定状态向量 x_k 的最优估计 $\hat{x}_k = \hat{x}_{k|k}$ 。可以得到以下结论。

定理 4.1 从数据 v_0, v_1, \dots, v_k 中确定的 x_k 的最优估计 $\hat{x}_k = \hat{x}_{k|k}$, 可以通过如下的递推方法进行计算。

定义:

$$P_{0,0} = \text{Var}(x_0)$$

对于 $k=1, 2, \dots$, 计算

$$\begin{aligned} P_{k,k-1} &= (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1}) P_{k-1,k-1} (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1})^T \\ &\quad + \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1} \mathbf{\Gamma}_{k-1}^T - \mathbf{K}_{k-1} \mathbf{R}_{k-1} \mathbf{K}_{k-1}^T \end{aligned} \quad (a)$$

其中

$$\mathbf{K}_{k-1} = \mathbf{\Gamma}_{k-1} \mathbf{S}_{k-1} \mathbf{R}_{k-1}^{-1} \quad (b)$$

则卡尔曼增益矩阵为

$$\mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \quad (c)$$

且

$$\mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \quad (d)$$

初始状态为

$$\hat{x}_{0|0} = E(x_0)$$

对于 $k=1, 2, \dots$, 预测估计为

$$\hat{x}_{k|k-1} = \mathbf{A}_{k-1} \hat{x}_{k-1|k-1} + \mathbf{K}_{k-1} (v_{k-1} - \mathbf{C}_{k-1} \hat{x}_{k-1|k-1}) \quad (e)$$

以及校正估计为

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \mathbf{G}_k (v_k - \mathbf{C}_k \hat{x}_{k|k-1}) \quad (f)$$

滤波算法结构框图见图 4.1。

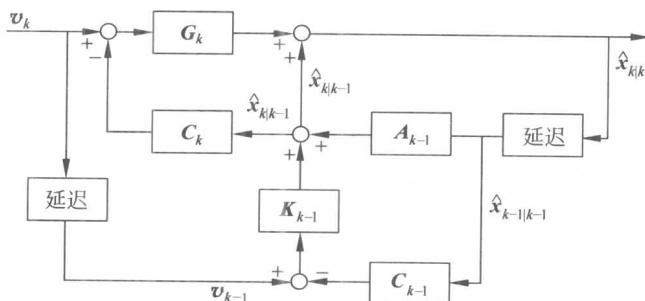


图 4.1 系统噪声和量测噪声相关的卡尔曼滤波算法结构框图

这就是系统噪声和量测噪声相关系统的卡尔曼滤波方程。如果系统噪声和量测噪声不相关, 即 $\mathbf{S}_{k-1} = \mathbf{O}_{p \times q}$, 则对于所有 $k=1, 2, \dots$, 都有 $\mathbf{K}_{k-1} = \mathbf{O}_{n \times q}$, 从而上面的卡尔曼滤波方程可以简化为第 2、3 章中讨论的情况。

我们首先推导预测—校正公式(e)和(f)。在这个过程中,将定义矩阵 $\mathbf{P}_{k,k-1}$ 、 $\mathbf{P}_{k,k}$ 和 \mathbf{G}_k ,并确定它们的计算程序(a)、(b)、(c)和(d)。令

$$\mathbf{v}^{k-1} = \begin{bmatrix} \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_{k-1} \end{bmatrix}, \quad \mathbf{v}^k = \begin{bmatrix} \mathbf{v}^{k-1} \\ \mathbf{v}_k \end{bmatrix}$$

则 \mathbf{v}^k 、 \mathbf{v}^{k-1} 和 \mathbf{v}_k 可以分别看作是引理4.2中的数据向量 \mathbf{v} 、 \mathbf{v}^1 和 \mathbf{v}^2 。同时,令

$$\hat{\mathbf{x}}_{k|k-1} = L(\mathbf{x}_k, \mathbf{v}^{k-1})$$

$$\hat{\mathbf{x}}_{k|k} = L(\mathbf{x}_k, \mathbf{v}^k)$$

以及

$$\mathbf{x}_k^\# = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} = \mathbf{x}_k - L(\mathbf{x}_k, \mathbf{v}^{k-1})$$

则有下列性质成立:

$$\left\{ \begin{array}{ll} \langle \underline{\xi}_{k-1}, \mathbf{v}^{k-2} \rangle = \mathbf{0}, & \langle \underline{\eta}_{k-1}, \mathbf{v}^{k-2} \rangle = \mathbf{0} \\ \langle \underline{\xi}_{k-1}, \mathbf{x}_{k-1} \rangle = \mathbf{0}, & \langle \underline{\eta}_{k-1}, \mathbf{x}_{k-1} \rangle = \mathbf{0} \\ \langle \mathbf{x}_{k-1}^\#, \underline{\xi}_{k-1} \rangle = \mathbf{0}, & \langle \mathbf{x}_{k-1}^\#, \underline{\eta}_{k-1} \rangle = \mathbf{0} \\ \langle \hat{\mathbf{x}}_{k-1|k-2}, \underline{\xi}_{k-1} \rangle = \mathbf{0}, & \langle \hat{\mathbf{x}}_{k-1|k-2}, \underline{\eta}_{k-1} \rangle = \mathbf{0} \end{array} \right. \quad (4.17)$$

(见练习4.4)。

为了推导预测公式,我们的思路是在下面的估计中增加“零值项” $\mathbf{K}_{k-1}(\mathbf{v}_{k-1} - \mathbf{C}_{k-1}\mathbf{x}_{k-1} - \underline{\eta}_{k-1})$:

$$\hat{\mathbf{x}}_{k|k-1} = L(\mathbf{A}_{k-1}\mathbf{x}_{k-1} + \Gamma_{k-1}\underline{\xi}_{k-1}, \mathbf{v}^{k-1})$$

选择一个合适的矩阵 \mathbf{K}_{k-1} ,可以将估计中噪声相关性的影响在 \mathbf{K}_{k-1} 中考虑。更确切地说,由于 $L(\cdot, \mathbf{v}^{k-1})$ 是线性的,有:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= L(\mathbf{A}_{k-1}\mathbf{x}_{k-1} + \Gamma_{k-1}\underline{\xi}_{k-1} + \mathbf{K}_{k-1}(\mathbf{v}_{k-1} - \mathbf{C}_{k-1}\mathbf{x}_{k-1} - \underline{\eta}_{k-1}), \mathbf{v}^{k-1}) \\ &= L((\mathbf{A}_{k-1} - \mathbf{K}_{k-1}\mathbf{C}_{k-1})\mathbf{x}_{k-1} + \mathbf{K}_{k-1}\mathbf{v}_{k-1} + (\Gamma_{k-1}\underline{\xi}_{k-1} - \mathbf{K}_{k-1}\underline{\eta}_{k-1}), \mathbf{v}^{k-1}) \\ &= (\mathbf{A}_{k-1} - \mathbf{K}_{k-1}\mathbf{C}_{k-1})L(\mathbf{x}_{k-1}, \mathbf{v}^{k-1}) + \mathbf{K}_{k-1}L(\mathbf{v}_{k-1}, \mathbf{v}^{k-1}) \\ &\quad + L(\Gamma_{k-1}\underline{\xi}_{k-1} - \mathbf{K}_{k-1}\underline{\eta}_{k-1}, \mathbf{v}^{k-1}) \\ &:= \mathbf{I}_1 + \mathbf{I}_2 + \mathbf{I}_3 \end{aligned}$$

我们将通过选择合适的 \mathbf{K}_{k-1} 使噪声项 \mathbf{I}_3 成为零。由式(4.6)和式(4.17),可得

$$\begin{aligned} \mathbf{I}_3 &= L(\Gamma_{k-1}\underline{\xi}_{k-1} - \mathbf{K}_{k-1}\underline{\eta}_{k-1}, \mathbf{v}^{k-1}) \\ &= E(\Gamma_{k-1}\underline{\xi}_{k-1} - \mathbf{K}_{k-1}\underline{\eta}_{k-1}) \end{aligned}$$

$$\begin{aligned}
& + \langle \boldsymbol{\Gamma}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} - \boldsymbol{K}_{k-1} \underline{\boldsymbol{\eta}}_{k-1}, \mathbf{v}^{k-1} \rangle [\|\mathbf{v}^{k-1}\|^2]^{-1} (\mathbf{v}^{k-1} - E(\mathbf{v}^{k-1})) \\
& = \left\langle \boldsymbol{\Gamma}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} - \boldsymbol{K}_{k-1} \underline{\boldsymbol{\eta}}_{k-1}, \begin{bmatrix} \mathbf{v}^{k-2} \\ \mathbf{v}_{k-1} \end{bmatrix} \right\rangle [\|\mathbf{v}^{k-1}\|^2]^{-1} (\mathbf{v}^{k-1} - E(\mathbf{v}^{k-1})) \\
& \stackrel{*}{=} \langle \boldsymbol{\Gamma}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} - \boldsymbol{K}_{k-1} \underline{\boldsymbol{\eta}}_{k-1}, \mathbf{v}_{k-1} \rangle [\|\mathbf{v}^{k-1}\|^2]^{-1} (\mathbf{v}^{k-1} - E(\mathbf{v}^{k-1})) \\
& = \langle \boldsymbol{\Gamma}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} - \boldsymbol{K}_{k-1} \underline{\boldsymbol{\eta}}_{k-1}, \mathbf{C}_{k-1} \mathbf{x}_{k-1} + \underline{\boldsymbol{\eta}}_{k-1} \rangle [\|\mathbf{v}^{k-1}\|^2]^{-1} (\mathbf{v}^{k-1} - E(\mathbf{v}^{k-1})) \\
& = \langle \boldsymbol{\Gamma}_{k-1} \mathbf{S}_{k-1} - \boldsymbol{K}_{k-1} \mathbf{R}_{k-1}, \mathbf{v}^{k-1} \rangle [\|\mathbf{v}^{k-1}\|^2]^{-1} (\mathbf{v}^{k-1} - E(\mathbf{v}^{k-1}))
\end{aligned}$$

因此,通过选择

$$\boldsymbol{K}_{k-1} = \boldsymbol{\Gamma}_{k-1} \mathbf{S}_{k-1} \mathbf{R}_{k-1}^{-1}$$

可以满足式(b),得到 $\mathbf{I}_3 = \mathbf{0}$ 。

其次, \mathbf{I}_1 和 \mathbf{I}_2 由下式决定:

$$\begin{aligned}
\mathbf{I}_1 &= (\mathbf{A}_{k-1} - \boldsymbol{K}_{k-1} \mathbf{C}_{k-1}) L(\mathbf{x}_{k-1}, \mathbf{v}^{k-1}) \\
&= (\mathbf{A}_{k-1} - \boldsymbol{K}_{k-1} \mathbf{C}_{k-1}) \hat{\mathbf{x}}_{k-1|k-1}
\end{aligned}$$

由引理 4.2 以及 $\mathbf{v}_{k-1}^\# = \mathbf{v}_{k-1} - L(\mathbf{v}_{k-1}, \mathbf{v}^{k-2})$, 可得

$$\begin{aligned}
\mathbf{I}_2 &= \boldsymbol{K}_{k-1} L(\mathbf{v}_{k-1}, \mathbf{v}^{k-1}) \\
&= \boldsymbol{K}_{k-1} \mathbf{L} \begin{pmatrix} \mathbf{v}_{k-1}, \begin{bmatrix} \mathbf{v}^{k-2} \\ \mathbf{v}_{k-1} \end{bmatrix} \end{pmatrix} \\
&= \boldsymbol{K}_{k-1} (L(\mathbf{v}_{k-1}, \mathbf{v}^{k-2}) + \langle \mathbf{v}_{k-1}^\#, \mathbf{v}_{k-1}^\# \rangle [\|\mathbf{v}_{k-1}^\#\|^2]^{-1} \mathbf{v}_{k-1}^\#) \\
&= \boldsymbol{K}_{k-1} (L(\mathbf{v}_{k-1}, \mathbf{v}^{k-2}) + \mathbf{v}_{k-1}^\#) \\
&= \boldsymbol{K}_{k-1} \mathbf{v}_{k-1}
\end{aligned}$$

因此,可以得到

$$\begin{aligned}
\hat{\mathbf{x}}_{k|k-1} &= \mathbf{I}_1 + \mathbf{I}_2 \\
&= (\mathbf{A}_{k-1} - \boldsymbol{K}_{k-1} \mathbf{C}_{k-1}) \hat{\mathbf{x}}_{k-1|k-1} + \boldsymbol{K}_{k-1} \mathbf{v}_{k-1} \\
&= \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \boldsymbol{K}_{k-1} (\mathbf{v}_{k-1} - \mathbf{C}_{k-1} \hat{\mathbf{x}}_{k-1|k-1})
\end{aligned}$$

上式即是预测公式(e)。

为推导校正公式,再次应用引理 4.2, 得

$$\begin{aligned}
\hat{\mathbf{x}}_{k|k} &= L(\mathbf{x}_k, \mathbf{v}^{k-1}) + \langle \mathbf{x}_k^\#, \mathbf{v}_k^\# \rangle [\|\mathbf{v}_k^\#\|^2]^{-1} \mathbf{v}_k^\# \\
&= \hat{\mathbf{x}}_{k|k-1} + \langle \mathbf{x}_k^\#, \mathbf{v}_k^\# \rangle [\|\mathbf{v}_k^\#\|^2]^{-1} \mathbf{v}_k^\#
\end{aligned} \tag{4.18}$$

式中

$$\mathbf{x}_k^\# = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}$$

由式(4.6)和式(4.17),可得

$$\mathbf{v}_k^\# = \mathbf{v}_k - L(\mathbf{v}_k, \mathbf{v}^{k-1})$$

$$\begin{aligned}
&= \mathbf{C}_k \mathbf{x}_k + \underline{\eta}_k - L(\mathbf{C}_k \mathbf{x}_k + \underline{\eta}_k, \mathbf{v}^{k-1}) \\
&= \mathbf{C}_k \mathbf{x}_k + \underline{\eta}_k - \mathbf{C}_k L(\mathbf{x}_k, \mathbf{v}^{k-1}) - L(\underline{\eta}_k, \mathbf{v}^{k-1}) \\
&= \mathbf{C}_k (\mathbf{x}_k - L(\mathbf{x}_k, \mathbf{v}^{k-1})) + \underline{\eta}_k - E(\underline{\eta}_k) \\
&\quad - \langle \underline{\eta}_k, \mathbf{v}^{k-1} \rangle [\parallel \mathbf{v}^{k-1} \parallel^2]^{-1} (\mathbf{v}^{k-1} - E(\mathbf{v}^{k-1})) \\
&= \mathbf{C}_k (\mathbf{x}_k - L(\mathbf{x}_k, \mathbf{v}^{k-1})) + \underline{\eta}_k \\
&= \mathbf{C}_k (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \underline{\eta}_k
\end{aligned}$$

再次应用式(4.17),由式(4.18)可知

$$\begin{aligned}
\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \langle \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}, \mathbf{C}_k (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \underline{\eta}_k \rangle \\
&\quad \cdot [\parallel \mathbf{C}_k (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \underline{\eta}_k \parallel^2_q]^{-1} (\mathbf{C}_k (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \underline{\eta}_k) \\
&= \hat{\mathbf{x}}_{k|k-1} + \parallel \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \parallel_n^2 \mathbf{C}_k^T \\
&\quad \cdot [\mathbf{C}_k \parallel \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \parallel_n^2 \mathbf{C}_k^T + \mathbf{R}_k]^{-1} (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \\
&= \hat{\mathbf{x}}_{k|k-1} + G_k (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})
\end{aligned}$$

上式即是校正公式(f),其中

$$\mathbf{P}_{k,j} = \parallel \mathbf{x}_k - \hat{\mathbf{x}}_{k|j} \parallel_n^2$$

且

$$\mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \quad (4.19)$$

剩下的工作就是验证关于 $\mathbf{P}_{k,k-1}$ 和 $\mathbf{P}_{k,k}$ 的递推关系(a)和(d)。推导过程需要应用下面两个公式,其证明由读者完成:

$$(\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \mathbf{C}_k = \mathbf{G}_k \mathbf{R}_k \quad (4.20)$$

及

$$\langle \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{\Gamma}_{k-1} \underline{\xi}_{k-1} - \mathbf{K}_{k-1} \underline{\eta}_{k-1} \rangle = \mathbf{O}_{n \times n} \quad (4.21)$$

(见练习 4.5)。

现在顺次应用式(e)、(b)和式(4.21),可得

$$\begin{aligned}
&\mathbf{P}_{k,k-1} \\
&= \parallel \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \parallel_n^2 \\
&= \parallel \mathbf{A}_{k-1} \mathbf{x}_{k-1} + \mathbf{\Gamma}_{k-1} \underline{\xi}_{k-1} - \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} - \mathbf{K}_{k-1} (\mathbf{v}_{k-1} - \mathbf{C}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}) \parallel_n^2 \\
&= \parallel \mathbf{A}_{k-1} \mathbf{x}_{k-1} + \mathbf{\Gamma}_{k-1} \underline{\xi}_{k-1} - \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} - \mathbf{K}_{k-1} (\mathbf{C}_{k-1} \mathbf{x}_{k-1} + \underline{\eta}_{k-1} - \mathbf{C}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}) \parallel_n^2 \\
&= \parallel (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1}) (\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}) + (\mathbf{\Gamma}_{k-1} \underline{\xi}_{k-1} - \mathbf{K}_{k-1} \underline{\eta}_{k-1}) \parallel_n^2
\end{aligned}$$

$$\begin{aligned}
&= (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1}) \mathbf{P}_{k-1|k-1} (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1})^T + \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1} \mathbf{\Gamma}_{k-1}^T \\
&\quad + \mathbf{K}_{k-1} \mathbf{R}_{k-1} \mathbf{K}_{k-1}^T - \mathbf{\Gamma}_{k-1} \mathbf{S}_{k-1} \mathbf{K}_{k-1}^T - \mathbf{K}_{k-1} \mathbf{S}_{k-1}^T \mathbf{\Gamma}_{k-1}^T \\
&= (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1}) \mathbf{P}_{k-1,k-1} (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1})^T \\
&\quad + \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1} \mathbf{\Gamma}_{k-1}^T - \mathbf{K}_{k-1} \mathbf{R}_{k-1} \mathbf{K}_{k-1}^T
\end{aligned}$$

这就是式(a)。

最后顺次应用式(f)、式(4.17)和式(4.20),可得

$$\begin{aligned}
\mathbf{P}_{k,k} &= \| \mathbf{x}_k - \hat{\mathbf{x}}_{k|k} \|^2_n \\
&= \| \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \|^2_n \\
&= \| (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) - \mathbf{G}_k (\mathbf{C}_k \mathbf{x}_k + \underline{\boldsymbol{\eta}}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \|^2_n \\
&= \| (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) - \mathbf{G}_k \underline{\boldsymbol{\eta}}_k \|^2_n \\
&= (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k)^T + \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^T \\
&\stackrel{*}{=} (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} - (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \mathbf{C}_k^T \mathbf{G}_k^T + \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^T \\
&= (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1}
\end{aligned}$$

上式是式(d)。

至此,定理全部证明完毕。

4.5 实时应用

飞机雷达制导系统为卡尔曼滤波的一个典型应用。该系统可以用第3章讨论的模型(3.26)来描述,只需做一个修改,即这里由机载跟踪雷达提供位置数据信息。因为系统和测量噪声过程来自相同的干扰源(如振动),它们是相关的。考虑以下的状态空间描述

$$\left\{ \begin{array}{l} \begin{bmatrix} x_{k+1}[1] \\ x_{k+1}[2] \\ x_{k+1}[3] \end{bmatrix} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_k[1] \\ x_k[2] \\ x_k[3] \end{bmatrix} + \begin{bmatrix} \xi_k[1] \\ \xi_k[2] \\ \xi_k[3] \end{bmatrix} \\ v_k = [1 \ 0 \ 0] \begin{bmatrix} x_k[1] \\ x_k[2] \\ x_k[3] \end{bmatrix} + \eta_k \end{array} \right. \quad (4.22)$$

式中 $\xi_k := [\xi_k[1] \ \xi_k[2] \ \xi_k[3]]^T$ 与 $\{\eta_k\}$ 假定是相关的零均值高斯白噪声序列,满足:

$$E(\xi_k) = \mathbf{0}, \quad E(\eta_k) = 0$$

$$E(\underline{\xi}\underline{\xi}^T) = \underline{Q}_k \delta_{kl}, \quad e(\eta_k \eta_l) = r_k \delta_{kl}, \quad E(\underline{\xi}_k \eta_l) = s_k \delta_{kl}$$

$$E(\underline{x}_0 \underline{\xi}_k^T) = \mathbf{0}, \quad E(\underline{x}_0 \eta_k) = \mathbf{0}$$

式中对于所有的 k , 都有 $\underline{Q}_k \geqslant \mathbf{0}$, $r_k > 0$, $s_k := [s_k[1] \ s_k[2] \ s_k[3]]^T \geqslant \mathbf{0}$, 且 $E(\underline{x}_0)$ 和 $\text{Var}(\underline{x}_0)$ 假定都是已知的。

对该系统应用定理 4.1, 可得

$$\underline{\mathbf{P}}_{k,k-1}[1,1] = \underline{\mathbf{P}}_{k-1}[1,1] + 2h\underline{\mathbf{P}}_{k-1}[1,2] + h^2 \underline{\mathbf{P}}_{k-1}[1,3] + h^2 \underline{\mathbf{P}}_{k-1}[2,2]$$

$$+ h^3 \underline{\mathbf{P}}_{k-1}[2,3] + \frac{h^4}{4} \underline{\mathbf{P}}_{k-1}[3,3] + \underline{Q}_{k-1}[1,1]$$

$$+ \frac{s_{k-1}[1]}{r_{k-1}} \left\{ \frac{s_{k-1}[1]}{r_{k-1}} \underline{\mathbf{P}}_{k-1}[1,1] - 2(\underline{\mathbf{P}}_{k-1}[1,1]$$

$$+ h\underline{\mathbf{P}}_{k-1}[1,2] + \frac{h^2}{2} \underline{\mathbf{P}}_{k-1}[1,3]) - s_{k-1}[1] \right\}$$

$$\underline{\mathbf{P}}_{k,k-1}[1,2] = \underline{\mathbf{P}}_{k,k-1}[2,1]$$

$$= \underline{\mathbf{P}}_{k-1}[1,2] + h\underline{\mathbf{P}}_{k-1}[1,3] + h\underline{\mathbf{P}}_{k-1}[2,2] + \frac{3h^2}{2} \underline{\mathbf{P}}_{k-1}[2,3]$$

$$+ \frac{h^3}{2} \underline{\mathbf{P}}_{k-1}[3,3] + \underline{Q}_{k-1}[1,2] + \left\{ \frac{s_{k-1}[1]s_{k-1}[2]}{r_{k-1}^2} \underline{\mathbf{P}}_{k-1}[1,1]$$

$$- \frac{s_{k-1}[1]}{r_{k-1}} (\underline{\mathbf{P}}_{k-1}[1,2] + h\underline{\mathbf{P}}_{k-1}[1,3])$$

$$- \frac{s_{k-1}[2]}{r_{k-1}} (\underline{\mathbf{P}}_{k-1}[1,1] + h\underline{\mathbf{P}}_{k-1}[1,2] + \frac{h^2}{2} \underline{\mathbf{P}}_{k-1}[1,3]) - \frac{s_{k-1}[1]s_{k-1}[2]}{r_{k-1}} \right\}$$

$$\underline{\mathbf{P}}_{k,k-1}[2,2] = \underline{\mathbf{P}}_{k-1}[2,2] + 2h\underline{\mathbf{P}}_{k-1}[2,3] + h^2 \underline{\mathbf{P}}_{k-1}[3,3] + \underline{Q}_{k-1}[2,2]$$

$$+ \frac{s_{k-1}[2]}{r_{k-1}} \left\{ \frac{s_{k-1}[2]}{r_{k-1}} \underline{\mathbf{P}}_{k-1}[2,2] - 2(\underline{\mathbf{P}}_{k-1}[1,2] + h\underline{\mathbf{P}}_{k-1}[1,3]) - s_{k-1}[2] \right\}$$

$$\underline{\mathbf{P}}_{k,k-1}[1,3] = \underline{\mathbf{P}}_{k,k-1}[3,1]$$

$$= \underline{\mathbf{P}}_{k-1}[1,3] + h\underline{\mathbf{P}}_{k-1}[2,3] + \frac{h^2}{2} \underline{\mathbf{P}}_{k-1}[3,3] + \underline{Q}_{k-1}[1,3]$$

$$+ \left\{ \frac{s_{k-1}[1]s_{k-1}[3]}{r_{k-1}^2} \underline{\mathbf{P}}_{k-1}[1,1] - \frac{s_{k-1}[1]}{r_{k-1}} \underline{\mathbf{P}}_{k-1}[1,3] \right.$$

$$\left. - \frac{s_{k-1}[3]}{r_{k-1}} (\underline{\mathbf{P}}_{k-1}[1,1] + h\underline{\mathbf{P}}_{k-1}[1,2] + \frac{h^2}{2} \underline{\mathbf{P}}_{k-1}[1,3]) - \frac{s_{k-1}[1]s_{k-1}[3]}{r_{k-1}} \right\}$$

$$\underline{\mathbf{P}}_{k,k-1}[2,3] = \underline{\mathbf{P}}_{k,k-1}[3,2]$$

$$= \underline{\mathbf{P}}_{k-1}[2,3] + h\underline{\mathbf{P}}_{k-1}[3,3] + \underline{Q}_{k-1}[2,3] + \left\{ \frac{s_{k-1}[2]s_{k-1}[3]}{r_{k-1}^2} \underline{\mathbf{P}}_{k-1}[1,1]$$

$$- \frac{s_{k-1}[2}s_{k-1}[3]}{r_{k-1}} - \frac{s_{k-1}[3]}{r_{k-1}} (\underline{\mathbf{P}}_{k-1}[1,2] + h\underline{\mathbf{P}}_{k-1}[1,3]) - \frac{s_{k-1}[2]s_{k-1}[3]}{r_{k-1}} \right\}$$

$$\begin{aligned}\mathbf{P}_{k,k-1}[3,3] &= \mathbf{P}_{k-1}[3,3] + \mathbf{Q}_{k-1}[3,3] \\ &\quad + \frac{s_{k-1}[3]}{r_{k-1}} \left\{ \frac{s_{k-1}[3]}{r_{k-1}} \mathbf{P}_{k-1}[1,1] - 2\mathbf{P}_{k-1}[1,3] - s_{k-1}[3] \right\}\end{aligned}$$

式中 $\mathbf{P}_{k-1} = \mathbf{P}_{k,k-1}$, $\mathbf{P}[i,j]$ 表示 \mathbf{P} 的第 (i,j) 个元素。

此外,

$$\begin{aligned}\mathbf{G}_k &= \frac{1}{\mathbf{P}_{k,k-1}[1,1] + r_k} \begin{bmatrix} \mathbf{P}_{k,k-1}[1,1] \\ \mathbf{P}_{k,k-1}[1,2] \\ \mathbf{P}_{k,k-1}[1,3] \end{bmatrix} \\ \mathbf{P}_k &= \mathbf{P}_{k,k-1} - \frac{1}{\mathbf{P}_{k,k-1}[1,1] + r_k} \cdot \begin{bmatrix} \mathbf{P}_{k,k-1}^2[1,1] & \mathbf{P}_{k,k-1}[1,1]\mathbf{P}_{k,k-1}[1,2] & \mathbf{P}_{k,k-1}[1,1]\mathbf{P}_{k,k-1}[1,3] \\ \mathbf{P}_{k,k-1}[1,1]\mathbf{P}_{k,k-1}[1,2] & \mathbf{P}_{k,k-1}^2[1,2] & \mathbf{P}_{k,k-1}[1,2]\mathbf{P}_{k,k-1}[1,3] \\ \mathbf{P}_{k,k-1}[1,1]\mathbf{P}_{k,k-1}[1,3] & \mathbf{P}_{k,k-1}[1,2]\mathbf{P}_{k,k-1}[1,3] & \mathbf{P}_{k,k-1}^2[1,3] \end{bmatrix}\end{aligned}$$

式中 $\mathbf{P}_0 = \text{Var}(\mathbf{x}_0)$, 且

$$\begin{aligned}\begin{bmatrix} \hat{\mathbf{x}}_{k|k-1}[1] \\ \hat{\mathbf{x}}_{k|k-1}[2] \\ \hat{\mathbf{x}}_{k|k-1}[3] \end{bmatrix} &= \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k-1|k-1}[1] \\ \hat{\mathbf{x}}_{k-1|k-1}[2] \\ \hat{\mathbf{x}}_{k-1|k-1}[3] \end{bmatrix} + \frac{v_k - \hat{\mathbf{x}}_{k-1|k-1}[1]}{r_{k-1}} \begin{bmatrix} s_{k-1}[1] \\ s_{k-1}[2] \\ s_{k-1}[3] \end{bmatrix} \\ \begin{bmatrix} \hat{\mathbf{x}}_{k|k}[1] \\ \hat{\mathbf{x}}_{k|k}[2] \\ \hat{\mathbf{x}}_{k|k}[3] \end{bmatrix} &= \begin{bmatrix} 1 - G_k[1] & (1 - G_k[1])h & (1 - G_k[1])h^2/2 \\ -G_k[2] & 1 - hG_k[2] & h - h^2G_k[2]/2 \\ -G_k[3] & -hG_k[3] & 1 - h^2G_k[3]/2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k-1|k-1}[1] \\ \hat{\mathbf{x}}_{k-1|k-1}[2] \\ \hat{\mathbf{x}}_{k-1|k-1}[3] \end{bmatrix} \\ &\quad + \begin{bmatrix} G_k[1] \\ G_k[2] \\ G_k[3] \end{bmatrix} v_k\end{aligned}$$

式中, $\hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0)$ 。

4.6 线性确定/随机系统

最后, 我们讨论包含确定性控制输入 u_k 的一般线性随机系统。更准确地说, 有以下状态空间描述:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k u_k + \boldsymbol{\Gamma}_k \underline{\boldsymbol{\xi}}_k \\ \mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k u_k + \underline{\boldsymbol{\eta}}_k \end{cases} \quad (4.23)$$

式中 u_k 是一个 m 维的确定性控制输入向量, $1 \leq m \leq n$ 。可以证明(见练习 4.6)应用于该系统的卡尔曼滤波算法由下式给出:

$$\left\{ \begin{array}{l} \mathbf{P}_{0,0} = \text{Var}(\mathbf{x}_0) \\ \mathbf{K}_{k-1} = \mathbf{\Gamma}_{k-1} \mathbf{S}_{k-1} \mathbf{R}_{k-1}^{-1} \\ \mathbf{P}_{k,k-1} = (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1}) \mathbf{P}_{k-1,k-1} (\mathbf{A}_{k-1} - \mathbf{K}_{k-1} \mathbf{C}_{k-1})^T \\ \quad + \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1} \mathbf{\Gamma}_{k-1}^T - \mathbf{K}_{k-1} \mathbf{R}_{k-1} \mathbf{K}_{k-1}^T \\ \mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ \mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \\ \hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \\ \hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_{k-1} u_{k-1} \\ \quad + \mathbf{K}_{k-1} (\mathbf{v}_{k-1} - \mathbf{D}_{k-1} u_{k-1} - \mathbf{C}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}) \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{D}_k u_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \\ k = 1, 2, \dots \end{array} \right. \quad (4.24)$$

(见图 4.2)。

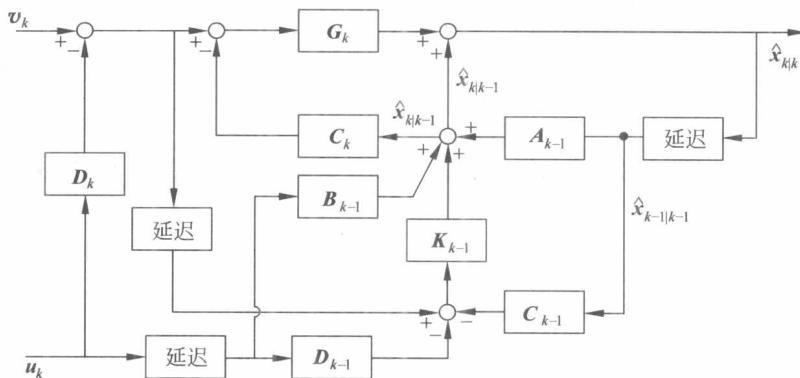


图 4.2 含确定输入且噪声相关时的卡尔曼滤波算法

如果系统和测量噪声过程是不相关的, 则对于所有的 $k=0, 1, 2, \dots$, 都有 $\mathbf{S}_k=\mathbf{0}$, 从而如我们所预料的那样, 系统(4.24)简化为系统(2.18)或系统(3.25)。

练习

4.1 假定 v 是一个随机向量, 并且定义: $L(\mathbf{x}, \mathbf{v}) = E(\mathbf{x}) + \langle \mathbf{x}, \mathbf{v} \rangle [\|\mathbf{v}\|^2]^{-1}(\mathbf{v} - E(\mathbf{v}))$ 。对于任意的常数矩阵 \mathbf{A}, \mathbf{B} 和随机向量 \mathbf{x}, \mathbf{y} , 证明

$L(\cdot, v)$ 是下式意义下的一个线性算子:

$$L(Ax + By, v) = AL(x, v) + BL(y, v)$$

4.2 假定 v 是一个随机向量, 并且 $L(\cdot, v)$ 的定义同上。证明如果 a 是一个常值向量, 则 $L(a, v) = a$ 。

4.3 对于一个实值函数 f 和一个矩阵 $A = [a_{ij}]$, 定义

$$\frac{df}{dA} = \left[\frac{\partial f}{\partial a_{ij}} \right]^T$$

取

$$\frac{\partial}{\partial H} (\text{tr } \|x - y\|^2) = \mathbf{0}$$

证明: 最小化问题

$$\text{tr } \|x^* - y\|^2 = \min_H \text{tr } \|x - y\|^2$$

的解 x^* (式中 $y = E(x) + H(v - E(v))$) 可以通过令

$$x^* = E(x) - \langle x, v \rangle [\|v\|^2]^{-1} (E(v) - v)$$

来获得, 并且 $H^* = \langle x, v \rangle [\|v\|^2]^{-1}$ 。

4.4 考虑线性随机系统(4.16)。令

$$v^{k-1} = \begin{bmatrix} v_0 \\ \vdots \\ v_{k-1} \end{bmatrix}, \quad v^k = \begin{bmatrix} v^{k-1} \\ v_k \end{bmatrix}$$

如练习 4.1 中一样定义 $L(x, v)$, 令 $x_k^\# = x_k - \hat{x}_{k|k-1}$, 其中 $\hat{x}_{k|k-1} := L(x_k, v^{k-1})$ 。

证明:

$$\langle \underline{\xi}_{k-1}, v^{k-2} \rangle = \mathbf{0}, \quad \langle \underline{\eta}_{k-1}, v^{k-2} \rangle = \mathbf{0}$$

$$\langle \underline{\xi}_{k-1}, x_{k-1} \rangle = \mathbf{0}, \quad \langle \underline{\eta}_{k-1}, x_{k-1} \rangle = \mathbf{0}$$

$$\langle x_{k-1}^\#, \underline{\xi}_{k-1} \rangle = \mathbf{0}, \quad \langle x_{k-1}^\#, \underline{\eta}_{k-1} \rangle = \mathbf{0}$$

$$\langle \hat{x}_{k-1|k-2}, \underline{\xi}_{k-1} \rangle = \mathbf{0}, \quad \langle \hat{x}_{k-1|k-2}, \underline{\eta}_{k-1} \rangle = \mathbf{0}$$

4.5 证明:

$$(I - G_k C_k) P_{k,k-1} C_k = G_k R_k$$

和

$$\langle x_{k-1} - \hat{x}_{k-1|k-1}, \Gamma_{k-1} \underline{\xi}_{k-1} - K_{k-1} \underline{\eta}_{k-1} \rangle = O_{n \times n}$$

4.6 考虑线性确定/随机系统:

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k + \Gamma_k \underline{\xi}_k \\ v_k = C_k x_k + D_k u_k + \underline{\eta}_k \end{cases}$$

式中 $\{u_k\}$ 是一个给定的确定性控制输入序列。假定它同样满足系统(4.16)的假设,推导该模型的卡尔曼滤波算法。

4.7 考虑下列在信号处理中的 ARMAX 模型(外源性自回归滑动平均模型):

$$v_k = -a_1 v_{k-1} - a_2 v_{k-2} - a_3 v_{k-3} + b_0 u_k + b_1 u_{k-1} + b_2 u_{k-2} + c_0 e_k + c_1 e_{k-1}$$

式中 $\{v_j\}$ 和 $\{u_j\}$ 分别表示输出和输入信号, $\{e_j\}$ 是一个零均值高斯白噪声序列, 且 $\text{Var}(e_j) = s_j > 0$, a_j 、 b_j 和 c_j 都是常值。

(a) 推导该 ARMAX 模型的一个状态空间描述;

(b) 推导该状态空间描述的卡尔曼滤波算法。

4.8 更加一般地,考虑在信号处理中的 ARMAX 模型:

$$v_k = - \sum_{j=1}^n a_j v_{k-j} + \sum_{j=0}^m b_j u_{k-j} + \sum_{j=0}^l c_j e_{k-j}$$

式中, $0 \leq m, l \leq n$, $\{v_j\}$ 和 $\{u_j\}$ 分别表示输出和输入信号, $\{e_j\}$ 是一个零均值高斯白噪声序列, 且 $\text{Var}(e_j) = s_j > 0$, a_j 、 b_j 和 c_j 都是常值。

(a) 推导该 ARMAX 模型的一个状态空间描述;

(b) 推导该状态空间描述的卡尔曼滤波算法。

第 5 章

有色噪声

考虑下面的状态空间描述的线性随机系统

$$\begin{cases} \underline{\mathbf{x}}_{k+1} = \mathbf{A}_k \underline{\mathbf{x}}_k + \underline{\boldsymbol{\Gamma}}_k \underline{\boldsymbol{\xi}}_k \\ \underline{\mathbf{v}}_k = \mathbf{C}_k \underline{\mathbf{x}}_k + \underline{\boldsymbol{\eta}}_k \end{cases} \quad (5.1)$$

式中 \mathbf{A}_k 、 $\underline{\boldsymbol{\Gamma}}_k$ 和 \mathbf{C}_k 分别是 $n \times n$ 、 $n \times p$ 和 $q \times n$ 阶已知常值矩阵, 且 $1 \leqslant p, q \leqslant n$ 。需要在下面的假设下, 基于初始条件 $E(\underline{\mathbf{x}}_0)$ 和 $\text{Var}(\underline{\mathbf{x}}_0)$, 给出 $\underline{\mathbf{x}}_k$ 的线性无偏最小方差估计:

$$(i) \quad \underline{\boldsymbol{\xi}}_k = \mathbf{M}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} + \underline{\boldsymbol{\beta}}_k$$

$$(ii) \quad \underline{\boldsymbol{\eta}}_k = \mathbf{N}_{k-1} \underline{\boldsymbol{\eta}}_{k-1} + \underline{\boldsymbol{\gamma}}_k$$

式中 $\underline{\boldsymbol{\xi}}_{-1} = \underline{\boldsymbol{\eta}}_{-1} = \mathbf{0}$, $\{\underline{\boldsymbol{\beta}}_k\}$ 和 $\{\underline{\boldsymbol{\gamma}}_k\}$ 是不相关的零均值高斯白噪声序列, 满足

$$E(\underline{\boldsymbol{\beta}}_k \underline{\boldsymbol{\gamma}}_l^T) = \mathbf{0}, \quad E(\underline{\boldsymbol{\beta}}_k \underline{\boldsymbol{\beta}}_l^T) = \mathbf{Q}_k \delta_{kl}, \quad E(\underline{\boldsymbol{\gamma}}_k \underline{\boldsymbol{\gamma}}_l^T) = \mathbf{R}_k \delta_{kl}$$

\mathbf{M}_{k-1} 和 \mathbf{N}_{k-1} 为已知的 $p \times p$ 和 $q \times q$ 阶常值矩阵。若噪声序列 $\{\underline{\boldsymbol{\xi}}_k\}$ 和 $\{\underline{\boldsymbol{\eta}}_k\}$ 满足条件(i)和(ii), 就称为有色噪声过程。本章主要研究噪声序列满足该假设的卡尔曼滤波。

5.1 处理思路

处理有色噪声模型(5.1)的思路, 是首先使得式(5.1)的系统方程白化。为了实现这个目标, 简单地假设

$$\mathbf{z}_k = \begin{bmatrix} \mathbf{x}_k \\ \boldsymbol{\xi}_k \end{bmatrix}, \quad \tilde{\mathbf{A}}_k = \begin{bmatrix} \mathbf{A}_k & \boldsymbol{\Gamma}_k \\ \mathbf{0} & \mathbf{M}_k \end{bmatrix}, \quad \tilde{\boldsymbol{\beta}}_k = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\beta}_k \end{bmatrix}$$

可得

$$\mathbf{z}_{k+1} = \tilde{\mathbf{A}}_k \mathbf{z}_k + \tilde{\boldsymbol{\beta}}_{k+1} \quad (5.2)$$

式(5.1)的量测方程变为

$$\mathbf{v}_k = \tilde{\mathbf{C}}_k \mathbf{z}_k + \boldsymbol{\eta}_k \quad (5.3)$$

式中 $\tilde{\mathbf{C}}_k = [\mathbf{C}_k \quad \mathbf{0}]$ 。

我们使用与第4章同样的模型

$$\hat{\mathbf{z}}_{k|j} = L(\mathbf{z}_k, \mathbf{v}^j)$$

式中

$$\mathbf{v}^j = \begin{bmatrix} \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_j \end{bmatrix}, \quad \hat{\mathbf{z}}_k := \hat{\mathbf{z}}_{k|k}$$

其次,推导 $\hat{\mathbf{z}}_k$ 的递推公式(代替预测-校正关系)。为了实现该步骤,利用 L 线性特性,可得

$$\hat{\mathbf{z}}_k = \tilde{\mathbf{A}}_{k-1} L(\mathbf{z}_{k-1}, \mathbf{v}^k) + L(\tilde{\boldsymbol{\beta}}_k, \mathbf{v}^k)$$

从噪声的假设可得

$$L(\tilde{\boldsymbol{\beta}}_k, \mathbf{v}^k) = \mathbf{0} \quad (5.4)$$

故

$$\hat{\mathbf{z}}_k = \tilde{\mathbf{A}}_{k-1} \hat{\mathbf{z}}_{k-1|k} \quad (5.5)$$

(见练习 5.1)。

为得到 $\hat{\mathbf{z}}_k$ 的递推公式,需将 $\hat{\mathbf{z}}_{k-1|k}$ 用 $\hat{\mathbf{z}}_{k-1}$ 表示。这可以通过引理 4.2 完成:

$$\begin{aligned} \hat{\mathbf{z}}_{k-1|k} &= L\left(\mathbf{z}_{k-1}, \begin{bmatrix} \mathbf{v}^{k-1} \\ \mathbf{v}_k \end{bmatrix}\right) \\ &= \hat{\mathbf{z}}_{k-1} + \langle \mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}, \mathbf{v}_k^\# \rangle [\|\mathbf{v}_k^\#\|^2]^{-1} \mathbf{v}_k^\# \end{aligned} \quad (5.6)$$

式中 $\hat{\mathbf{z}}_{k-1} = L(\mathbf{z}_{k-1}, \mathbf{v}^{k-1})$ 和 $\mathbf{v}_k^\# = \mathbf{v}_k - L(\mathbf{v}_k, \mathbf{v}^{k-1})$ 。

5.2 误差估计

现在最重要的是理解式(5.6)中的误差项。首先推导 $\mathbf{v}_k^\#$ 的公式。由 $\mathbf{v}_k^\#$ 的定义和观测方程式(5.3),及

$$\tilde{\mathbf{C}}_k \tilde{\boldsymbol{\beta}}_k = [\mathbf{C}_k \quad \mathbf{0}] \begin{bmatrix} \mathbf{0} \\ \tilde{\boldsymbol{\beta}}_k \end{bmatrix} = \mathbf{0}$$

可得

$$\begin{aligned} \mathbf{v}_k &= \tilde{\mathbf{C}}_k \mathbf{z}_k + \underline{\boldsymbol{\eta}}_k \\ &= \tilde{\mathbf{C}}_k \mathbf{z}_k + \mathbf{N}_{k-1} \underline{\boldsymbol{\eta}}_{k-1} + \underline{\boldsymbol{\gamma}}_k \\ &= \tilde{\mathbf{C}}_k (\tilde{\mathbf{A}}_{k-1} \mathbf{z}_{k-1} + \tilde{\boldsymbol{\beta}}_k) + \mathbf{N}_{k-1} (\mathbf{v}_{k-1} - \tilde{\mathbf{C}}_{k-1} \mathbf{z}_{k-1}) + \underline{\boldsymbol{\gamma}}_k \\ &= \mathbf{H}_{k-1} \mathbf{z}_{k-1} + \mathbf{N}_{k-1} \mathbf{v}_{k-1} + \underline{\boldsymbol{\gamma}}_k \end{aligned} \quad (5.7)$$

且

$$\begin{aligned} \mathbf{H}_{k-1} &= \tilde{\mathbf{C}}_k \tilde{\mathbf{A}}_{k-1} - \mathbf{N}_{k-1} \tilde{\mathbf{C}}_{k-1} \\ &= [\mathbf{C}_k \mathbf{A}_{k-1} - \mathbf{N}_{k-1} \mathbf{C}_{k-1} \quad \mathbf{C}_k \boldsymbol{\Gamma}_{k-1}] \end{aligned}$$

利用 L 的线性特性, 可得

$$L(\mathbf{v}_k, \mathbf{v}^{k-1}) = \mathbf{H}_{k-1} L(\mathbf{z}_{k-1}, \mathbf{v}^{k-1}) + \mathbf{N}_{k-1} L(\mathbf{v}_{k-1}, \mathbf{v}^{k-1}) + L(\underline{\boldsymbol{\gamma}}_k, \mathbf{v}^{k-1})$$

因为 $L(\mathbf{z}_{k-1}, \mathbf{v}^{k-1}) = \hat{\mathbf{z}}_{k-1}$, 则

$$L(\mathbf{v}_{k-1}, \mathbf{v}^{k-1}) = \mathbf{v}^{k-1} \quad (5.8)$$

(见练习 5.2), 以及

$$L(\underline{\boldsymbol{\gamma}}_{k-1}, \mathbf{v}^{k-1}) = \mathbf{0} \quad (5.9)$$

可得

$$\begin{aligned} \mathbf{v}_k^\# &= \mathbf{v}_k - L(\mathbf{v}_k, \mathbf{v}^{k-1}) \\ &= \mathbf{v}_k - (\mathbf{H}_{k-1} \hat{\mathbf{z}}_{k-1} + \mathbf{N}_{k-1} \mathbf{v}_{k-1}) \\ &= \mathbf{v}_k - \mathbf{N}_{k-1} \mathbf{v}_{k-1} - \mathbf{H}_{k-1} \hat{\mathbf{z}}_{k-1} \end{aligned} \quad (5.10)$$

此外, 由式(5.7)和式(5.10), 可得

$$\mathbf{v}_k^\# = \mathbf{H}_{k-1} (\mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}) + \underline{\boldsymbol{\gamma}}_k \quad (5.11)$$

对式(5.6), 根据式(5.11)、式(5.10)及 $\langle \mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}, \underline{\boldsymbol{\gamma}}_k \rangle = \mathbf{0}$ (见练习 5.3), 可得

$$\begin{aligned} \hat{\mathbf{z}}_{k-1|k} &= \hat{\mathbf{z}}_{k-1} + \langle \mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}, \mathbf{v}_k^\# \rangle [\|\mathbf{v}_k^\#\|^2]^{-1} \mathbf{v}_k^\# \\ &= \hat{\mathbf{z}}_{k-1} + \|\mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}\|^2 \mathbf{H}_{k-1}^T (\mathbf{H}_{k-1} \|\mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}\|^2 \mathbf{H}_{k-1}^T + \mathbf{R}_k)^{-1} \\ &\quad \cdot (\mathbf{v}_k - \mathbf{N}_{k-1} \mathbf{v}_{k-1} - \mathbf{H}_{k-1} \hat{\mathbf{z}}_{k-1}) \end{aligned}$$

代入式(5.5), 有

$$\hat{\mathbf{z}}_k = \tilde{\mathbf{A}}_{k-1} \hat{\mathbf{z}}_{k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{N}_{k-1} \mathbf{v}_{k-1} - \mathbf{H}_{k-1} \hat{\mathbf{z}}_{k-1})$$

即

$$\begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{\boldsymbol{\xi}}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{k-1} & \boldsymbol{\Gamma}_{k-1} \\ \mathbf{0} & \mathbf{M}_{k-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \hat{\boldsymbol{\xi}}_{k-1} \end{bmatrix} + \mathbf{G}_k \left(\mathbf{v}_k - \mathbf{N}_{k-1} \mathbf{v}_{k-1} - \mathbf{H}_{k-1} \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \hat{\boldsymbol{\xi}}_{k-1} \end{bmatrix} \right) \quad (5.12)$$

式中

$$\begin{aligned} \mathbf{G}_k &= \tilde{\mathbf{A}}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T (\mathbf{H}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T + \mathbf{R}_k)^{-1} \\ &= \begin{bmatrix} \mathbf{A}_{k-1} & \boldsymbol{\Gamma}_{k-1} \\ \mathbf{0} & \mathbf{M}_{k-1} \end{bmatrix} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T (\mathbf{H}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T + \mathbf{R}_k)^{-1} \end{aligned} \quad (5.13)$$

且

$$\mathbf{P}_k = \| \mathbf{z}_k - \hat{\mathbf{z}}_k \|^2 = \text{Var}(\mathbf{z}_k - \hat{\mathbf{z}}_k) \quad (5.14)$$

5.3 卡尔曼滤波过程

剩下的工作就是推导 \mathbf{P}_k 的算法和初始条件 $\hat{\mathbf{z}}_0$ 。

应用式(5.7), 有

$$\begin{aligned} \mathbf{z}_k - \hat{\mathbf{z}}_k &= (\tilde{\mathbf{A}}_{k-1} \mathbf{z}_{k-1} + \tilde{\boldsymbol{\beta}}_k) - (\tilde{\mathbf{A}}_{k-1} \hat{\mathbf{z}}_{k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{N}_{k-1} \mathbf{v}_{k-1} - \mathbf{H}_{k-1} \hat{\mathbf{z}}_{k-1})) \\ &= \tilde{\mathbf{A}}_{k-1} (\mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}) + \tilde{\boldsymbol{\beta}}_k - \mathbf{G}_k (\mathbf{H}_{k-1} \mathbf{z}_{k-1} + \boldsymbol{\gamma}_k - \mathbf{H}_{k-1} \hat{\mathbf{z}}_{k-1}) \\ &= (\tilde{\mathbf{A}}_{k-1} - \mathbf{G}_k \mathbf{H}_{k-1}) (\mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}) + (\tilde{\boldsymbol{\beta}}_k - \mathbf{G}_k \boldsymbol{\gamma}_k) \end{aligned}$$

此外, 从练习 5.3 和 $\langle \mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}, \tilde{\boldsymbol{\beta}}_k \rangle = \mathbf{0}$ (见练习 5.4), 有

$$\begin{aligned} \mathbf{P}_k &= (\tilde{\mathbf{A}}_{k-1} - \mathbf{G}_k \mathbf{H}_{k-1}) \mathbf{P}_{k-1} (\tilde{\mathbf{A}}_{k-1} - \mathbf{G}_k \mathbf{H}_{k-1})^T + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & Q_k \end{bmatrix} + \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^T \\ &= (\tilde{\mathbf{A}}_{k-1} - \mathbf{G}_k \mathbf{H}_{k-1}) \mathbf{P}_{k-1} \tilde{\mathbf{A}}_{k-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & Q_k \end{bmatrix} \end{aligned} \quad (5.15)$$

上面使用了根据式(5.13)得来的下面的恒等式:

$$\begin{aligned} &- (\tilde{\mathbf{A}}_{k-1} - \mathbf{G}_k \mathbf{H}_{k-1}) \mathbf{P}_{k-1} (\mathbf{G}_k \mathbf{H}_{k-1})^T + \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^T \\ &= -\tilde{\mathbf{A}}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T \mathbf{G}_k^T + \mathbf{G}_k \mathbf{H}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T \mathbf{G}_k^T + \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^T \\ &= -\tilde{\mathbf{A}}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T \mathbf{G}_k^T + \mathbf{G}_k (\mathbf{H}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T + \mathbf{R}_k) \mathbf{G}_k^T \\ &= -\tilde{\mathbf{A}}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T \mathbf{G}_k^T + \tilde{\mathbf{A}}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T \mathbf{G}_k^T \\ &= \mathbf{0} \end{aligned}$$

初始估计为

$$\hat{\mathbf{z}}_0 = L(\mathbf{z}_0, \mathbf{v}_0) = E(\mathbf{z}_0) - \langle \mathbf{z}_0, \mathbf{v}_0 \rangle [\| \mathbf{v}_0 \|^2]^{-1} (E(\mathbf{v}_0) - \mathbf{v}_0)$$

$$= \begin{bmatrix} E(\mathbf{x}_0) \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T \\ \mathbf{0} \end{bmatrix} [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T + \mathbf{R}_0]^{-1} (\mathbf{C}_0 E(\mathbf{x}_0) - \mathbf{v}_0) \quad (5.16)$$

注意到

$$E(\hat{\mathbf{z}}_0) = \begin{bmatrix} E(\mathbf{x}_0) \\ \mathbf{0} \end{bmatrix} = E(\mathbf{z}_0)$$

并且因为

$$\begin{aligned} \mathbf{z}_k - \hat{\mathbf{z}}_k &= (\tilde{\mathbf{A}}_{k-1} - \mathbf{G}_k \mathbf{H}_{k-1})(\mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}) + (\tilde{\boldsymbol{\beta}}_k - \mathbf{G}_k \boldsymbol{\gamma}_k) \\ &= (\tilde{\mathbf{A}}_{k-1} - \mathbf{G}_k \mathbf{H}_{k-1}) \cdots (\tilde{\mathbf{A}}_0 - \mathbf{G}_1 \mathbf{H}_0)(\mathbf{z}_0 - \hat{\mathbf{z}}_0) + \text{噪声} \end{aligned}$$

从而有 $E(\mathbf{z}_k - \hat{\mathbf{z}}_k) = \mathbf{0}$, 也就是, $\hat{\mathbf{z}}_k$ 是 \mathbf{z}_k 的线性无偏最小(误差)方差估计。

此外,由式(5.16),可得初始条件:

$$\begin{aligned} P_0 &= \text{Var}(\mathbf{z}_0 - \hat{\mathbf{z}}_0) \\ &= \text{Var} \left(\begin{bmatrix} \mathbf{x}_0 - E(\mathbf{x}_0) \\ \boldsymbol{\xi}_0 \end{bmatrix} + \begin{bmatrix} \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T \\ \mathbf{0} \end{bmatrix} \cdot [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T + \mathbf{R}_0]^{-1} [\mathbf{C}_0 E(\mathbf{x}_0) - \mathbf{v}_0] \right) \\ &= \begin{bmatrix} \text{Var}(\mathbf{x}_0) + \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T \\ \cdot [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T + \mathbf{R}_0]^{-1} [\mathbf{C}_0 E(\mathbf{x}_0) - \mathbf{v}_0] \\ \mathbf{0} \end{bmatrix} \mathbf{Q}_0 \\ &= \begin{bmatrix} \text{Var}(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \\ \cdot [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T + \mathbf{R}_0]^{-1} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \\ \mathbf{0} \end{bmatrix} \mathbf{Q}_0 \quad (5.17a) \end{aligned}$$

(见练习 5.5)。若 $\text{Var}(\mathbf{x}_0)$ 非奇异,则根据矩阵求逆引理(见引理 1.2),有

$$\mathbf{P}_0 = \begin{bmatrix} ([\text{Var}(\mathbf{x}_0)]^{-1} + \mathbf{C}_0^T \mathbf{R}_0^{-1} \mathbf{C}_0)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix} \quad (5.17b)$$

最后,回到定义 $\mathbf{z}_k = \begin{bmatrix} \mathbf{x}_k \\ \boldsymbol{\xi}_k \end{bmatrix}$,有

$$\begin{aligned} \hat{\mathbf{z}}_k &= L(\mathbf{z}_k, \mathbf{v}^k) \\ &= E(\mathbf{z}_k) + \langle \mathbf{z}_k, \mathbf{v}^k \rangle [\|\mathbf{v}^k\|^2]^{-1} (\mathbf{v}^k - E(\mathbf{v}^k)) \\ &= \begin{bmatrix} E(\mathbf{x}_k) \\ E(\boldsymbol{\xi}_k) \end{bmatrix} + \begin{bmatrix} \langle \mathbf{x}_k, \mathbf{v}^k \rangle \\ \langle \boldsymbol{\xi}_k, \mathbf{v}^k \rangle \end{bmatrix} [\|\mathbf{v}^k\|^2]^{-1} (\mathbf{v}^k - E(\mathbf{v}^k)) \end{aligned}$$

$$= \begin{bmatrix} E(\mathbf{x}_k) + \langle \mathbf{x}_k, \mathbf{v}^k \rangle [\|\mathbf{v}^k\|^2]^{-1} (\mathbf{v}^k - E(\mathbf{v}^k)) \\ E(\hat{\boldsymbol{\xi}}_k) + \langle \hat{\boldsymbol{\xi}}_k, \mathbf{v}^k \rangle [\|\mathbf{v}^k\|^2]^{-1} (\mathbf{v}^k - E(\mathbf{v}^k)) \end{bmatrix}$$

$$= \begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{\boldsymbol{\xi}}_k \end{bmatrix}$$

总之,有色噪声模型(5.1)的卡尔曼滤波过程为

$$\begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{\boldsymbol{\xi}}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{k-1} & \boldsymbol{\Gamma}_{k-1} \\ \mathbf{0} & \mathbf{M}_{k-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \hat{\boldsymbol{\xi}}_{k-1} \end{bmatrix} + \mathbf{G}_k \left(\mathbf{v}_k - \mathbf{N}_{k-1} \mathbf{v}_{k-1} - \mathbf{H}_{k-1} \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \hat{\boldsymbol{\xi}}_{k-1} \end{bmatrix} \right) \quad (5.18)$$

式中

$$\mathbf{H}_{k-1} = [\mathbf{C}_k \mathbf{A}_{k-1} - \mathbf{N}_{k-1} \mathbf{C}_{k-1} \quad \mathbf{C}_k \boldsymbol{\Gamma}_{k-1}]$$

及

$$\mathbf{G}_k = \begin{bmatrix} \mathbf{A}_{k-1} & \boldsymbol{\Gamma}_{k-1} \\ \mathbf{0} & \mathbf{M}_{k-1} \end{bmatrix} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T (\mathbf{H}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T + \mathbf{R}_k)^{-1} \quad (5.19)$$

其中 \mathbf{P}_{k-1} 由下式可得:

$$\mathbf{P}_k = \left(\begin{bmatrix} \mathbf{A}_{k-1} & \boldsymbol{\Gamma}_{k-1} \\ \mathbf{0} & \mathbf{M}_{k-1} \end{bmatrix} - \mathbf{G}_k \mathbf{H}_{k-1} \right) \mathbf{P}_{k-1} \begin{bmatrix} \mathbf{A}_{k-1}^T & \mathbf{0} \\ \boldsymbol{\Gamma}_{k-1}^T & \mathbf{M}_{k-1}^T \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_k \end{bmatrix} \quad (5.20)$$

$k=1, 2, \dots$ 。初始条件由式(5.16)、(5.17a)或(5.17b)给出。

如果有色噪声过程变为白噪声(即对所有的 k , $\mathbf{M}_k = \mathbf{0}$ 和 $\mathbf{N}_k = \mathbf{0}$), 则此时的卡尔曼滤波算法简化为第2章和第3章推导的卡尔曼滤波算法, 这只需通过简单地假设

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}$$

使得 $\hat{\mathbf{x}}_k$ 的递推关系可以分解为两个方程: 预测和校正方程。另外, 通过定义 $\mathbf{P}_k = \mathbf{P}_{k|k}$ 及

$$\mathbf{P}_{k,k-1} = \mathbf{A}_{k-1} \mathbf{P}_{k-1} \mathbf{A}_{k-1}^T + \boldsymbol{\Gamma}_k \mathbf{Q}_k \boldsymbol{\Gamma}_k^T$$

可以看出, 式(5.20)就是计算 $\mathbf{P}_{k,k-1}$ 和 $\mathbf{P}_{k,k}$ 的算法。我们将这个推导作为练习留给读者解答(见练习5.6)。

5.4 系统白噪声

如果只有量测噪声是有色的, 而系统噪声是白噪声, 即 $\mathbf{M}_k = \mathbf{0}$, $\mathbf{N}_k \neq \mathbf{0}$, 则在推导卡尔曼滤波方程时, 没有必要计算额外的估计量 $\hat{\boldsymbol{\xi}}_k$ 。这种情况下, 滤波算法为(见练习5.7)

$$\left\{ \begin{array}{l} \mathbf{P}_0 = ([\text{Var}(\mathbf{x}_0)]^{-1} + \mathbf{C}_0^T \mathbf{R}_0^{-1} \mathbf{C}_0)^{-1} \\ \mathbf{H}_{k-1} = [\mathbf{C}_k \mathbf{A}_{k-1} - \mathbf{N}_{k-1} \mathbf{C}_{k-1}] \\ \mathbf{G}_k = (\mathbf{A}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T \mathbf{C}_k^T) \\ \quad \cdot (\mathbf{H}_{k-1} \mathbf{P}_{k-1} \mathbf{H}_{k-1}^T + \mathbf{C}_k \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T \mathbf{C}_k^T + \mathbf{R}_{k-1})^{-1} \\ \mathbf{P}_k = (\mathbf{A}_{k-1} - \mathbf{G}_k \mathbf{H}_{k-1}) \mathbf{P}_{k-1} \mathbf{A}_{k-1}^T + (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T \\ \hat{\mathbf{x}}_0 = E(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T + \mathbf{R}_0]^{-1} [\mathbf{C}_0 E(\mathbf{x}_0) - \mathbf{v}_0] \\ \hat{\mathbf{x}}_k = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{N}_{k-1} \mathbf{v}_{k-1} - \mathbf{H}_{k-1} \hat{\mathbf{x}}_{k-1}) \\ k = 1, 2, \dots \end{array} \right. \quad (5.21)$$

(见图 5.1)。

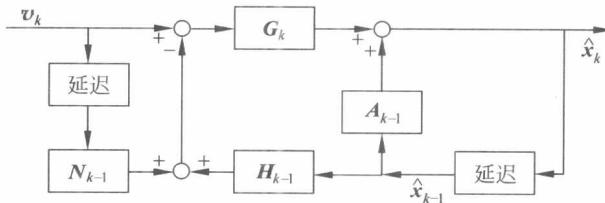


图 5.1 有色量测噪声时的滤波算法*

5.5 实时应用

考虑有色输入的跟踪系统(见练习 3.8 和式(3.26))。其状态空间描述为

$$\left\{ \begin{array}{l} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \underline{\xi}_k \\ \mathbf{v}_k = \mathbf{C}\mathbf{x}_k + \eta_k \end{array} \right. \quad (5.22)$$

式中

$$\mathbf{A} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = [1 \ 0 \ 0]$$

采样时间 $h>0$,且

$$\left\{ \begin{array}{l} \underline{\xi}_k = \mathbf{F}\underline{\xi}_{k-1} + \beta_k \\ \eta_k = g\eta_{k-1} + \gamma_k \end{array} \right. \quad (5.23)$$

* 译者注：原著图中为 \bar{x}_{k-1} , 疑有误, 已改为 \hat{x}_{k-1} 。

式中 $\{\underline{\beta}_k\}$ 和 $\{\gamma_k\}$ 都是零均值高斯白噪声序列,且满足以下假设:

$$E(\underline{\beta}_k \underline{\beta}_l^T) = Q_k \delta_{kl}, \quad E(\eta_k \eta_l) = r_k \delta_{kl}, \quad E(\underline{\beta}_k \gamma_l) = \mathbf{0}$$

$$E(\mathbf{x}_0 \underline{\xi}_0^T) = \mathbf{0}, \quad E(\mathbf{x}_0 \eta_0) = \mathbf{0}, \quad E(\mathbf{x}_0 \underline{\beta}_k^T) = \mathbf{0}$$

$$E(\mathbf{x}_0 \gamma_k) = \mathbf{0}, \quad E(\underline{\xi}_0 \underline{\beta}_k^T) = \mathbf{0}, \quad E(\underline{\xi}_0 \gamma_k) = \mathbf{0}$$

$$E(\eta_0 \underline{\beta}_k^T) = \mathbf{0}, \quad \underline{\xi}_{-1} = \mathbf{0}, \quad \eta_{-1} = 0$$

设

$$\mathbf{z}_k = \begin{bmatrix} \mathbf{x}_k \\ \underline{\xi}_k \end{bmatrix}, \quad \tilde{\mathbf{A}}_k = \begin{bmatrix} \mathbf{A}_k & \mathbf{I} \\ \mathbf{0} & \mathbf{F} \end{bmatrix}, \quad \tilde{\underline{\beta}}_k = \begin{bmatrix} \mathbf{0} \\ \underline{\beta}_k \end{bmatrix}, \quad \tilde{\mathbf{C}}_k = [\mathbf{C}_k \quad \mathbf{0}]$$

有

$$\begin{cases} \mathbf{z}_{k+1} = \tilde{\mathbf{A}}_k \mathbf{z}_k + \tilde{\underline{\beta}}_{k+1} \\ v_k = \tilde{\mathbf{C}}_k \mathbf{z}_k + \eta_k \end{cases}$$

这正如式(5.2)和式(5.3)所描述的。

相应的卡尔曼滤波算法由式(5.18)~(5.20),得到下面的形式:

$$\begin{cases} \begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{\underline{\xi}}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{F} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \hat{\underline{\xi}}_{k-1} \end{bmatrix} + \mathbf{G}_k \begin{bmatrix} v_k - g v_{k-1} - \mathbf{h}_{k-1}^T \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \hat{\underline{\xi}}_{k-1} \end{bmatrix} \end{bmatrix} \\ \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \hat{\underline{\xi}}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{E}(\mathbf{x}_0) \\ \mathbf{0} \end{bmatrix} \end{cases}$$

式中

$$\mathbf{h}_{k-1} = [\mathbf{CA} - g\mathbf{C} \quad \mathbf{C}]^T = [(1-g) \quad h \quad h^2/2 \quad 1 \quad 0 \quad 0]^T$$

$$\mathbf{G}_k = \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{F} \end{bmatrix} \mathbf{P}_{k-1} \mathbf{h}_{k-1} (\mathbf{h}_{k-1}^T \mathbf{P}_{k-1} \mathbf{h}_{k-1} + \boldsymbol{\gamma}_k)^{-1}$$

其中 \mathbf{P}_{k-1} 由下式给出:

$$\begin{cases} \mathbf{P}_k = \left(\begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{F} \end{bmatrix} - \mathbf{G}_k \mathbf{h}_{k-1}^T \right) \mathbf{P}_{k-1} \begin{bmatrix} \mathbf{A}^T & \mathbf{0} \\ \mathbf{I} & \mathbf{F}^T \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & Q_k \end{bmatrix} \\ \mathbf{P}_0 = \begin{bmatrix} \text{Var}(\mathbf{x}_0) & \mathbf{0} \\ \mathbf{0} & \text{Var}(\underline{\xi}_0) \end{bmatrix} \end{cases}$$

练习

5.1 令 $\{\underline{\beta}_k\}$ 是零均值高斯白噪声, $\{v_k\}$ 是与系统(5.1)中同样的观测数据序列, 并设

$$\tilde{\underline{\beta}}_k = \begin{bmatrix} \mathbf{0} \\ \underline{\beta}_k \end{bmatrix}, \quad v^k = \begin{bmatrix} v_0 \\ \vdots \\ v_k \end{bmatrix}$$

与式(4.6)一样定义 $L(x, v)$ 。证明: $L(\tilde{\underline{\beta}}_k, v^k) = \mathbf{0}$ 。

5.2 设 $\{\gamma_k\}$ 是零均值高斯白噪声序列, v^k 和 $L(x, v)$ 与上面的定义一致。证明:

$$\begin{aligned} L(v_{k-1}, v^{k-1}) &= v_{k-1} \\ L(\underline{\gamma}_k, v^{k-1}) &= \mathbf{0} \end{aligned}$$

5.3 设 $\{\gamma_k\}$ 是零均值高斯白噪声序列, v^k 和 $L(x, v)$ 与练习5.1的定义一致。此外, 设 $\hat{z}_{k-1} = L(z_{k-1}, v^{k-1})$, $z_{k-1} = \begin{bmatrix} x_{k-1} \\ \xi_{k-1} \end{bmatrix}$ 。证明: $\langle z_{k-1} - \hat{z}_{k-1}, \underline{\gamma}_k \rangle = \mathbf{0}$ 。

5.4 令 $\{\underline{\beta}_k\}$ 是零均值高斯白噪声序列, 并设 $\tilde{\underline{\beta}}_k = \begin{bmatrix} \mathbf{0} \\ \underline{\beta}_k \end{bmatrix}$, 与练习5.3一样定义 \hat{z}_{k-1} 。证明: $\langle z_{k-1} - \hat{z}_{k-1}, \tilde{\underline{\beta}}_k \rangle = \mathbf{0}$ 。

5.5 与式(4.6)一样定义 $L(x, v)$, 并设 $\hat{z}_0 = L(z_0, v_0)$, 且 $z_0 = \begin{bmatrix} x_0 \\ \xi_0 \end{bmatrix}$ 。

证明:

$$\text{Var}(z_0 - \hat{z}_0) = \begin{bmatrix} \text{Var}(x_0) - [\text{Var}(x_0)]C_0^T \\ \cdot [C_0 \text{Var}(x_0)C_0^T + R_0]^{-1} C_0 [\text{Var}(x_0)] \\ \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ Q_0 \end{bmatrix}.$$

5.6 证明: 如果系统(5.1)中的矩阵 M_k 和 N_k 对所有的 k 都为零, 则式(5.18)~(5.20)给出的卡尔曼滤波算法简化为第2章和第3章推导出的系统噪声和量测噪声为不相关白噪声的线性随机系统的卡尔曼滤波算法。

5.7 当 $M_k = \mathbf{0}$ 但 $N_k \neq \mathbf{0}$ 时, 简化系统(5.1)的卡尔曼滤波算法。

5.8 考虑以(5.23)的有色噪声为输入的跟踪系统(5.22)，

(a) 设

$$\underline{\mathbf{X}}_k = \begin{bmatrix} \underline{x}_k \\ \underline{\xi}_k \\ \eta_k \end{bmatrix}, \quad \underline{\zeta}_k = \begin{bmatrix} \mathbf{0} \\ \underline{\beta}_{k+1} \\ \gamma_{k+1} \end{bmatrix}$$

$$\mathbf{A}_c = \begin{bmatrix} \mathbf{A} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & g \end{bmatrix}, \quad \mathbf{C}_c = [\mathbf{C} \ 0 \ 0 \ 0 \ 1]$$

将该有色噪声输入的系统重新整理为高斯白噪声输入的增广系统；

(b) 通过应用公式(3.25)于该增广系统，给出具有有色输入(5.23)的跟踪系统(5.22)的卡尔曼滤波算法；

(c) 该方案的主要缺点是什么？

第 6 章

极限(稳态)卡尔曼滤波

本章,我们考虑所有系统矩阵与时间无关的特殊情形。也就是说,将要研究线性时不变随机系统,其状态空间可描述为

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \boldsymbol{\Gamma}\underline{\xi}_k \\ \mathbf{v}_k = \mathbf{C}\mathbf{x}_k + \underline{\eta}_k \end{cases} \quad (6.1)$$

这里 \mathbf{A} 、 $\boldsymbol{\Gamma}$ 和 \mathbf{C} 分别是 $n \times n$ 、 $n \times p$ 和 $q \times n$ 阶常值矩阵, $1 \leq p, q \leq n$, $\{\underline{\xi}_k\}$ 和 $\{\underline{\eta}_k\}$ 分别是零均值高斯白噪声序列,且

$$E(\underline{\xi}_k \underline{\xi}_l^\top) = \mathbf{Q}\delta_{kl}, \quad E(\underline{\eta}_k \underline{\eta}_l^\top) = \mathbf{R}\delta_{kl}, \quad E(\underline{\xi}_k \underline{\eta}_l^\top) = \mathbf{0}$$

式中 \mathbf{Q} 和 \mathbf{R} 分别是与 k 无关的 $p \times p$ 和 $q \times q$ 阶非负定和正定对称矩阵。

该特殊情形的卡尔曼滤波算法可以描述为(见图 6.1)

$$\begin{cases} \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1}) \\ \hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} \\ \hat{\mathbf{x}}_{0|0} = \mathbf{E}(\mathbf{x}_0) \end{cases} \quad (6.2)$$

且有

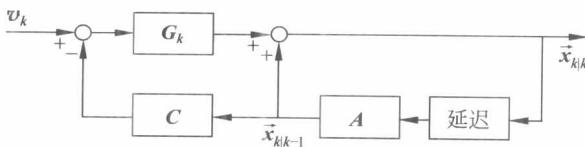


图 6.1 线性时不变系统的卡尔曼滤波算法

$$\begin{cases} \mathbf{P}_{0,0} = \text{Var}(\mathbf{x}_0) \\ \mathbf{P}_{k,k-1} = \mathbf{A}\mathbf{P}_{k-1,k-1}\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \\ \mathbf{G}_k = \mathbf{P}_{k,k-1}\mathbf{C}^T(\mathbf{C}\mathbf{P}_{k,k-1}\mathbf{C}^T + \mathbf{R})^{-1} \\ \mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k\mathbf{C})\mathbf{P}_{k,k-1} \end{cases} \quad (6.3)$$

注意到,即使对于这样的简单模型,也需要在完成预测-校正滤波(6.2)前,每次在计算(6.3)中卡尔曼增益矩阵 \mathbf{G}_k 时,进行矩阵求逆运算。在实时应用中,有时为了节约计算时间,需要用一个固定增益矩阵代替(6.2)中的卡尔曼增益矩阵 \mathbf{G}_k 。

用 $k \rightarrow \infty$ 时 \mathbf{G}_k 的极限 \mathbf{G} 代替 \mathbf{G}_k ,定义极限(或稳态)卡尔曼滤波,其中 \mathbf{G} 被称为极限卡尔曼增益矩阵。这时式(6.2)的预测-校正方程变为

$$\begin{cases} \vec{\mathbf{x}}_{k|k} = \vec{\mathbf{x}}_{k|k-1} + \mathbf{G}(\mathbf{v}_k - \mathbf{C}\vec{\mathbf{x}}_{k|k-1}) \\ \vec{\mathbf{x}}_{k|k-1} = \mathbf{A}\vec{\mathbf{x}}_{k-1|k-1} \\ \vec{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \end{cases} \quad (6.4)$$

在线性系统(6.1)变化非常平缓的条件下,可以证明序列 $\{\mathbf{G}_k\}$ 确实是收敛的。事实上, $\text{tr} \parallel \vec{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k} \parallel^2_n$ 以指数迅速趋于零。因此,用 \mathbf{G} 代替 \mathbf{G}_k 并不会太多地改变估计的最优化。

6.1 处理思路

根据(6.3)中 \mathbf{G}_k 的定义,为了研究 \mathbf{G}_k 的收敛性,有必要研究下式的收敛性:

$$\mathbf{P}_k := \mathbf{P}_{k,k-1}$$

首先建立 \mathbf{P}_k 的递推关系。因为

$$\begin{aligned} \mathbf{P}_k &= \mathbf{P}_{k,k-1} = \mathbf{A}\mathbf{P}_{k-1,k-1}\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \\ &= \mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})\mathbf{P}_{k-1,k-2}\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \\ &= \mathbf{A}(\mathbf{I} - \mathbf{P}_{k-1,k-2}\mathbf{C}^T(\mathbf{C}\mathbf{P}_{k-1,k-2}\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C})\mathbf{P}_{k-1,k-2}\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \\ &= \mathbf{A}(\mathbf{P}_{k-1} - \mathbf{P}_{k-1}\mathbf{C}^T(\mathbf{C}\mathbf{P}_{k-1}\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\mathbf{P}_{k-1})\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \end{aligned}$$

设

$$\boldsymbol{\Psi}(\mathbf{T}) = \mathbf{A}(\mathbf{T} - \mathbf{T}\mathbf{C}^T(\mathbf{C}\mathbf{T}\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\mathbf{T})\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T$$

则 \mathbf{P}_k 实际上满足递推关系:

$$\mathbf{P}_k = \boldsymbol{\Psi}(\mathbf{P}_{k-1}) \quad (6.5)$$

该关系称为矩阵 Riccati 方程。若当 $k \rightarrow \infty$ 时 $\mathbf{P}_k \rightarrow \mathbf{P}$, 则 \mathbf{P} 满足矩阵 Riccati 方程

$$\mathbf{P} = \boldsymbol{\Psi}(\mathbf{P}) \quad (6.6)$$

求解式(6.6)得到 \mathbf{P} ,并定义

$$\mathbf{G} = \mathbf{P}\mathbf{C}^T (\mathbf{C}\mathbf{P}\mathbf{C}^T + \mathbf{R})^{-1}$$

则 $k \rightarrow \infty$ 时, $\mathbf{G}_k \rightarrow \mathbf{G}$ 。并注意到 \mathbf{P}_k 是对称的, 故 $\Psi(\mathbf{P}_k)$ 也是对称的。

本章后面的论述将说明 $\{\mathbf{P}_k\}$ 实际上是收敛的, 只要下面的条件成立:

- (i) 对于所有的 k 存在常值对称矩阵 \mathbf{W} , 使得 $\mathbf{P}_k \leq \mathbf{W}$ (即对于所有的 k , $\mathbf{W} - \mathbf{P}_k$ 是非负定对称的);
- (ii) 对所有 $k=0, 1, \dots, P_k \leq P_{k+1}$;
- (iii) 存在常值对称矩阵 \mathbf{P} , 满足 $\lim_{k \rightarrow \infty} \mathbf{P}_k = \mathbf{P}$ 。

6.2 主要结论

为了获得唯一的 \mathbf{P} , 我们必须说明, 只要 $\mathbf{P}_0 \geq \mathbf{0}$, 它将是与初始条件 $\mathbf{P}_0 := \mathbf{P}_{0,-1}$ 无关的。

引理 6.1 假设线性系统(6.1)是可观测的, 即以下矩阵满秩:

$$\mathbf{N}_{CA} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix}$$

则存在独立于 \mathbf{P}_0 的非负定对称常值矩阵 \mathbf{W} , 使得对于所有 $k \geq n+1$, 有

$$\mathbf{P}_k \leq \mathbf{W}$$

由 $\langle \mathbf{x}_k - \hat{\mathbf{x}}_k, \underline{\xi}_k \rangle = 0$ (见式(3.20)), 可得

$$\begin{aligned} \mathbf{P}_k &:= \mathbf{P}_{k,k-1} = \| \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \|_n^2 \\ &= \| \mathbf{A}\mathbf{x}_{k-1} + \Gamma \underline{\xi}_{k-1} - \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} \|_n^2 \\ &= \mathbf{A} \| \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1} \|_n^2 \mathbf{A}^T + \Gamma \mathbf{Q} \Gamma^T \end{aligned}$$

同理, 因为 $\hat{\mathbf{x}}_{k-1|k-1}$ 是 \mathbf{x}_{k-1} 的最小方差估计, 对 \mathbf{x}_{k-1} 的其他任意无偏估计 $\tilde{\mathbf{x}}_{k-1}$, 有

$$\| \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1} \|_n^2 \leq \| \mathbf{x}_{k-1} - \tilde{\mathbf{x}}_{k-1} \|_n^2$$

根据 \mathbf{N}_{CA} 满秩的假设, 可知 $\mathbf{N}_{CA}^T \mathbf{N}_{CA}$ 非奇异。此外,

$$\mathbf{N}_{CA}^T \mathbf{N}_{CA} = \sum_{i=0}^{n-1} (\mathbf{A}^T)^i \mathbf{C}^T \mathbf{C} \mathbf{A}^i$$

按照下式来选择 $\tilde{\mathbf{x}}_{k-1}$:

$$\tilde{\mathbf{x}}_{k-1} = \mathbf{A}^n [\mathbf{N}_{CA}^T \mathbf{N}_{CA}]^{-1} \sum_{i=0}^{n-1} (\mathbf{A}^T)^i \mathbf{C}^T \mathbf{v}_{k-n-1+i}, \quad k \geq n+1 \quad (6.7)$$

很明显, $\tilde{\mathbf{x}}_{k-1}$ 关于数据是线性的, 还可以证明它是无偏的(见练习 6.1)。更

进一步,

$$\begin{aligned}\tilde{\mathbf{x}}_{k-1} &= \mathbf{A}^n [\mathbf{N}_{CA}^T \mathbf{N}_{CA}]^{-1} \sum_{i=0}^{n-1} (\mathbf{A}^T)^i \mathbf{C}^T (\mathbf{C} \mathbf{x}_{k-n-1+i} + \underline{\boldsymbol{\eta}}_{k-n-1+i}) \\ &= \mathbf{A}^n [\mathbf{N}_{CA}^T \mathbf{N}_{CA}]^{-1} \sum_{i=0}^{n-1} (\mathbf{A}^T)^i \mathbf{C}^T \left(\mathbf{C} \mathbf{A}^i \mathbf{x}_{k-n-1} + \sum_{j=0}^{i-1} \mathbf{C} \mathbf{A}^j \mathbf{T} \underline{\boldsymbol{\xi}}_{i-1-j} + \underline{\boldsymbol{\eta}}_{k-n-1+i} \right) \\ &= \mathbf{A}^n \mathbf{x}_{k-n-1} + \mathbf{A}^n [\mathbf{N}_{CA}^T \mathbf{N}_{CA}]^{-1} \sum_{i=0}^{n-1} (\mathbf{A}^T)^i \mathbf{C}^T \left(\sum_{j=0}^{i-1} \mathbf{C} \mathbf{A}^j \mathbf{T} \underline{\boldsymbol{\xi}}_{i-1-j} + \underline{\boldsymbol{\eta}}_{k-n-1+i} \right)\end{aligned}$$

又因

$$\mathbf{x}_{k-1} = \mathbf{A}^n \mathbf{x}_{k-n-1} + \sum_{i=0}^{n-1} \mathbf{A}^i \mathbf{T} \underline{\boldsymbol{\xi}}_{n-1+i}$$

可得

$$\begin{aligned}\mathbf{x}_{k-1} - \tilde{\mathbf{x}}_{k-1} &= \sum_{i=0}^{n-1} \mathbf{A}^i \mathbf{T} \underline{\boldsymbol{\xi}}_{n-1+i} \\ &\quad - \mathbf{A}^n [\mathbf{N}_{CA}^T \mathbf{N}_{CA}]^{-1} \sum_{i=0}^{n-1} (\mathbf{A}^T)^i \mathbf{C}^T \left(\sum_{j=0}^{i-1} \mathbf{C} \mathbf{A}^j \underline{\boldsymbol{\xi}}_{i-1-j} + \underline{\boldsymbol{\eta}}_{k-n-1+i} \right)\end{aligned}$$

注意到对所有 m 和 l 都有 $E(\underline{\boldsymbol{\xi}}_m \underline{\boldsymbol{\eta}}_l^T) = \mathbf{0}$ 和 $E(\underline{\boldsymbol{\eta}}_l \underline{\boldsymbol{\eta}}_l^T) = \mathbf{R}$, 故对所有 $k \geq n+1$,

$\|\mathbf{x}_{k-1} - \tilde{\mathbf{x}}_{k-1}\|_n^2$ 都与 k 无关。因此, 对所有 $k \geq n+1$, 有

$$\begin{aligned}\mathbf{P}_k &= \mathbf{A} \|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}\|_n^2 \mathbf{A}^T + \mathbf{T} \mathbf{Q} \mathbf{T}^T \\ &\leq \mathbf{A} \|\mathbf{x}_{k-1} - \tilde{\mathbf{x}}_{k-1|k-1}\|_n^2 \mathbf{A}^T + \mathbf{T} \mathbf{Q} \mathbf{T}^T \\ &= \mathbf{A} \|\mathbf{x}_n - \tilde{\mathbf{x}}_{n|n}\|_n^2 \mathbf{A}^T + \mathbf{T} \mathbf{Q} \mathbf{T}^T\end{aligned}$$

取

$$\mathbf{W} = \mathbf{A} \|\mathbf{x}_n - \tilde{\mathbf{x}}_{n|n}\|_n^2 \mathbf{A}^T + \mathbf{T} \mathbf{Q} \mathbf{T}^T$$

则对所有 $k \geq n+1$, 有 $\mathbf{P}_k \leq \mathbf{W}$ 。且 \mathbf{W} 与初始条件 $\mathbf{P}_0 = \mathbf{P}_{0,-1} = \|\mathbf{x}_0 - \hat{\mathbf{x}}_{0,-1}\|_n^2$ 无关。这就完成了引理的证明。

引理 6.2 若 \mathbf{P} 和 \mathbf{Q} 都是非负定对称矩阵, 且 $\mathbf{P} \geq \mathbf{Q}$, 则 $\Psi(\mathbf{P}) \geq \Psi(\mathbf{Q})$ 。 ■

为证明该引理, 引入公式

$$\frac{d}{ds} \mathbf{A}^{-1}(s) = -\mathbf{A}^{-1}(s) \left[\frac{d}{ds} \mathbf{A}(s) \right] \mathbf{A}^{-1}(s)$$

(见练习 6.2)。

记 $\mathbf{T}(s) = \mathbf{Q} + s(\mathbf{P} - \mathbf{Q})$, 有

$$\begin{aligned}\Psi(\mathbf{P}) - \Psi(\mathbf{Q}) &= \int_0^1 \frac{d}{ds} \Psi(\mathbf{Q} + s(\mathbf{P} - \mathbf{Q})) ds \\ &= \mathbf{A} \left\{ \int_0^1 \frac{d}{ds} \{(\mathbf{Q} + s(\mathbf{P} - \mathbf{Q})) - (\mathbf{Q} + s(\mathbf{P} - \mathbf{Q})) \mathbf{C}^T \right.\end{aligned}$$

$$\begin{aligned}
& \cdot [\mathbf{C}(\mathbf{Q} + s(\mathbf{P} - \mathbf{Q}))\mathbf{C}^T + \mathbf{R}]^{-1} \mathbf{C}(\mathbf{Q} + s(\mathbf{P} - \mathbf{Q})) \} ds \} \mathbf{A}^T \\
& = \mathbf{A} \left\{ \int_0^1 [\mathbf{P} - \mathbf{Q} - (\mathbf{P} - \mathbf{Q})\mathbf{C}^T (\mathbf{C}\mathbf{T}(s)\mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C}\mathbf{T}(s) \right. \\
& \quad \left. - \mathbf{T}(s)\mathbf{C}^T (\mathbf{C}\mathbf{T}(s)\mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C}(\mathbf{P} - \mathbf{Q}) + \mathbf{T}(s)\mathbf{C}^T (\mathbf{C}\mathbf{T}(s)\mathbf{C}^T + \mathbf{R})^{-1} \right. \\
& \quad \left. \cdot \mathbf{C}(\mathbf{P} - \mathbf{Q})\mathbf{C}^T (\mathbf{C}\mathbf{T}(s)\mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C}\mathbf{T}(s)] ds \right\} \mathbf{A}^T \\
& = \mathbf{A} \left\{ \int_0^1 [\mathbf{T}(s)\mathbf{C}^T (\mathbf{C}\mathbf{T}(s)\mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C}] (\mathbf{P} - \mathbf{Q}) \right. \\
& \quad \left. \cdot [\mathbf{T}(s)\mathbf{C}^T (\mathbf{C}\mathbf{T}(s)\mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C}]^T ds \right\} \mathbf{A} \\
& \geqslant 0
\end{aligned}$$

即 $\Psi(\mathbf{P}) \geqslant \Psi(\mathbf{Q})$ 。

还有下面的引理。

引理 6.3 假设线性系统(6.1)可观测, 则在初始条件为 $\mathbf{P}_0 = \mathbf{P}_{0,-1} = \mathbf{0}$ 时, 当 $k \rightarrow \infty$, 序列 $\{\mathbf{P}_k\}$ 收敛到某个对称矩阵 $\mathbf{P} \geqslant \mathbf{0}$ 。 ■

证明如下。

因为

$$\mathbf{P}_1 := \| \mathbf{x}_1 - \hat{\mathbf{x}}_{1|0} \|_n^2 \geqslant \mathbf{0} = \mathbf{P}_0$$

及 \mathbf{P}_0 和 \mathbf{P}_1 都是对称的, 根据引理 6.2 可得

$$\begin{aligned}
\mathbf{P}_2 &= \Psi(\mathbf{P}_1) \geqslant \Psi(\mathbf{P}_0) = \mathbf{P}_1 \\
&\dots \\
\mathbf{P}_{k+1} &\geqslant \mathbf{P}_k, \quad k = 0, 1, \dots
\end{aligned}$$

因此 $\{\mathbf{P}_k\}$ 是单调不减的, 并以 \mathbf{W} 为界(见引理 6.1)。

对任意 n 维向量 \mathbf{y} , 有

$$0 \leqslant \mathbf{y}^T \mathbf{P}_k \mathbf{y} \leqslant \mathbf{y}^T \mathbf{W} \mathbf{y}$$

所以序列 $\{\mathbf{y}^T \mathbf{P}_k \mathbf{y}\}$ 是有界非负单调不减的实数序列, 必定会收敛到某一非负实数。令

$$\mathbf{y} = [0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0]^T$$

第 i 个元素为 1。设 $\mathbf{P}_k = [\mathbf{p}_{ij}^{(k)}]$ 。存在非负数 p_{ii} , 当 $k \rightarrow \infty$ 时,

$$\mathbf{y}^T \mathbf{P}_k \mathbf{y} = p_{ii}^{(k)} \rightarrow p_{ii}$$

令

$$\mathbf{y} = [0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0]^T$$

其中第 i 和第 j 个元素为 1。则存在非负数 q , 当 $k \rightarrow \infty$ 时,

$$\begin{aligned}
\mathbf{y}^T \mathbf{P}_k \mathbf{y} &= p_{ii}^{(k)} + p_{ij}^{(k)} + p_{ji}^{(k)} + p_{jj}^{(k)} \\
&= p_{ii}^{(k)} + 2p_{ij}^{(k)} + p_{jj}^{(k)} \rightarrow q
\end{aligned}$$

因为 $p_{ii}^{(k)} \rightarrow p_{ii}$, 故当 $k \rightarrow \infty$ 时, $p_{ij}^{(k)} \rightarrow \frac{1}{2}(q - p_{ii} - p_{jj})$, 即 $\mathbf{P}_k \rightarrow \mathbf{P}$ 。

因为 $\mathbf{P}_k \geq 0$ 并且对称, 所以 \mathbf{P} 也一样。这就完成了引理的证明。

现在定义

$$\mathbf{G} = \lim_{k \rightarrow \infty} \mathbf{G}_k$$

式中 $\mathbf{G}_k = \mathbf{P}_k \mathbf{C}^T (\mathbf{C} \mathbf{P}_k \mathbf{C}^T + \mathbf{R})^{-1}$, 则

$$\mathbf{G} = \mathbf{P} \mathbf{C}^T (\mathbf{C} \mathbf{P} \mathbf{C}^T + \mathbf{R})^{-1} \quad (6.8)$$

下面将说明, 对于作为初始值的任意非负定对称矩阵 \mathbf{P}_0 , $\{\mathbf{P}_k\}$ 仍然收敛到同样的 \mathbf{P} 。因此引入一个任意的非负定对称矩阵 \mathbf{P}_0 , 定义 $\mathbf{P}_k = \Psi(\mathbf{P}_{k-1})$, $k=1, 2, \dots$ 和 $\mathbf{P} = \Psi(\mathbf{P})$ 。我们首先需要下面的结论。

引理 6.4 设线性系统(6.1)是可观测的, 使得 \mathbf{P} 可用引理 6.3 定义。则对所有的 $k=1, 2, \dots$, 和任意非负定对称初始条件 \mathbf{P}_0 , 有

$$\mathbf{P} - \mathbf{P}_k = \mathbf{A}(\mathbf{I} - \mathbf{G}\mathbf{C})(\mathbf{P} - \mathbf{P}_{k-1})(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})^T \mathbf{A}^T \quad (6.9) \blacksquare$$

证明如下。

由 $\mathbf{G}_{k-1} = \mathbf{P}_{k-1} \mathbf{C}^T (\mathbf{C} \mathbf{P}_{k-1} \mathbf{C}^T + \mathbf{R})^{-1}$ 和 $\mathbf{P}_{k-1}^T = \mathbf{P}_{k-1}$, 可知矩阵 $\mathbf{G}_{k-1} \mathbf{C} \mathbf{P}_{k-1}$ 是非负定对称的, 可得 $\mathbf{G}_{k-1} \mathbf{C} \mathbf{P}_{k-1} = \mathbf{P}_{k-1} \mathbf{C}^T \mathbf{G}_{k-1}^T$ 。根据式(6.5)和式(6.6), 有

$$\begin{aligned} \mathbf{P} - \mathbf{P}_k &= \Psi(\mathbf{P}) - \Psi(\mathbf{P}_{k-1}) \\ &= (\mathbf{A}\mathbf{P}\mathbf{A}^T - \mathbf{A}\mathbf{G}\mathbf{C}\mathbf{P}\mathbf{A}^T) - (\mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T - \mathbf{A}\mathbf{G}_{k-1}\mathbf{C}\mathbf{P}_{k-1}\mathbf{A}^T) \\ &= \mathbf{A}\mathbf{P}\mathbf{A}^T - \mathbf{A}\mathbf{G}\mathbf{C}\mathbf{P}\mathbf{A}^T - \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{A}\mathbf{P}_{k-1}\mathbf{C}^T\mathbf{G}_{k-1}^T\mathbf{A}^T \end{aligned} \quad (6.10)$$

现在

$$(\mathbf{I} - \mathbf{G}\mathbf{C})(\mathbf{P} - \mathbf{P}_{k-1})(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})^T = \mathbf{P} - \mathbf{P}_{k-1} + \mathbf{P}_{k-1}\mathbf{C}^T\mathbf{G}_{k-1}^T - \mathbf{G}\mathbf{C}\mathbf{P} + \mathbf{R}_e \quad (6.11)$$

式中

$$\mathbf{R}_e = \mathbf{G}\mathbf{C}\mathbf{P}_{k-1} - \mathbf{P}\mathbf{C}^T\mathbf{G}_{k-1}^T + \mathbf{G}\mathbf{C}\mathbf{P}\mathbf{C}^T\mathbf{G}_{k-1}^T - \mathbf{G}\mathbf{C}\mathbf{P}_{k-1}\mathbf{C}^T\mathbf{G}_{k-1}^T \quad (6.12)$$

因此, 如果能够说明 $\mathbf{R}_e = \mathbf{0}$, 则式(6.9)可从式(6.10)和式(6.11)推得。由 \mathbf{G}_{k-1} 的定义, 得 $\mathbf{G}_{k-1}(\mathbf{C} \mathbf{P}_{k-1} \mathbf{C}^T + \mathbf{R}) = \mathbf{P}_{k-1} \mathbf{C}^T$ 即 $(\mathbf{C} \mathbf{P}_{k-1} \mathbf{C}^T + \mathbf{R}) \mathbf{G}_{k-1} = \mathbf{C} \mathbf{P}_{k-1}$, 所以

$$\mathbf{G}_{k-1} \mathbf{C} \mathbf{P}_{k-1} \mathbf{C}^T = \mathbf{P}_{k-1} \mathbf{C}^T - \mathbf{G}_{k-1} \mathbf{R} \quad (6.13)$$

即

$$\mathbf{C} \mathbf{P}_{k-1} \mathbf{C}^T \mathbf{G}_{k-1}^T = \mathbf{C} \mathbf{P}_{k-1} - \mathbf{R} \mathbf{G}_{k-1}^T \quad (6.14)$$

式(6.13)中, 取初始条件 $\mathbf{P}_0 := \mathbf{P}_{0,-1} = \mathbf{0}$, 令 $k \rightarrow \infty$, 有

$$\mathbf{G} \mathbf{C} \mathbf{P} \mathbf{C}^T = \mathbf{P} \mathbf{C}^T - \mathbf{G} \mathbf{R} \quad (6.15)$$

将式(6.14)和式(6.15)代入式(6.12), 可得 $\mathbf{R}_e = \mathbf{0}$ 。

这就完成了引理 6.4 的证明。

引理 6.5

$$\mathbf{P}_k = [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]\mathbf{P}_{k-1}[\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]^T + [\mathbf{A}\mathbf{G}_{k-1}]\mathbf{R}[\mathbf{A}\mathbf{G}_{k-1}]^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \quad (6.16)$$

则对于初始条件为 $\mathbf{P}_0 := \mathbf{P}_{0,-1} = \mathbf{0}$ 的可观测系统, 有

$$\mathbf{P} = [\mathbf{A}(\mathbf{I} - \mathbf{G}\mathbf{C})]\mathbf{P}[\mathbf{A}(\mathbf{I} - \mathbf{G}\mathbf{C})]^T + [\mathbf{A}\mathbf{G}]\mathbf{R}[\mathbf{A}\mathbf{G}]^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \quad (6.17) \blacksquare$$

由定义有 $\mathbf{G}_{k-1}(\mathbf{C}\mathbf{P}_{k-1}\mathbf{C}^T + \mathbf{R}) = \mathbf{P}_{k-1}\mathbf{C}^T$, 从而有

$$\mathbf{G}_{k-1}\mathbf{R} = (\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})\mathbf{P}_{k-1}\mathbf{C}^T$$

因此

$$\mathbf{A}\mathbf{G}_{k-1}\mathbf{R}\mathbf{G}_{k-1}^T\mathbf{A}^T = \mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})\mathbf{P}_{k-1}\mathbf{C}^T\mathbf{G}_{k-1}^T\mathbf{A}^T$$

由矩阵 Riccati 方程 $\mathbf{P}_k = \boldsymbol{\Psi}(\mathbf{P}_{k-1})$, 可得

$$\begin{aligned} \mathbf{P}_k &= \mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})\mathbf{P}_{k-1}\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \\ &= \mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})\mathbf{P}_{k-1}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})^T\mathbf{A}^T + \mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})\mathbf{P}_{k-1}\mathbf{C}^T\mathbf{G}_{k-1}^T\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \\ &= \mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})\mathbf{P}_{k-1}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})^T\mathbf{A}^T + \mathbf{A}\mathbf{G}_{k-1}\mathbf{R}\mathbf{G}_{k-1}^T\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \end{aligned}$$

即式(6.16)。

引理 6.6 设线性系统(6.1)(完全)可控(即矩阵

$$\mathbf{M}_{\boldsymbol{\Gamma}} = [\boldsymbol{\Gamma} \quad \mathbf{A}\boldsymbol{\Gamma} \quad \cdots \quad \mathbf{A}^{n-1}\boldsymbol{\Gamma}]$$

满秩)。则对于任意非负定对称初始矩阵 \mathbf{P}_0 , 当 $k \geq n+1$ 时, 有 $\mathbf{P}_k > \mathbf{0}$, 所以, $\mathbf{P} > \mathbf{0}$ 。■

应用式(6.16) k 次, 有

$$\begin{aligned} \mathbf{P}_k &= \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T + [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]\boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T[\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]^T + \cdots \\ &\quad + \{[\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_2\mathbf{C})]\}\boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T\{[\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_2\mathbf{C})]\}^T \\ &\quad + [\mathbf{A}\mathbf{G}_{k-1}]\mathbf{R}[\mathbf{A}\mathbf{G}_{k-1}]^T + [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})][\mathbf{A}\mathbf{G}_{k-2}]\mathbf{R}[\mathbf{A}\mathbf{G}_{k-2}]^T \\ &\quad \cdot [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]^T + \cdots + \{[\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \\ &\quad \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_2\mathbf{C})][\mathbf{A}\mathbf{G}_1]\}\mathbf{R}\{[\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \\ &\quad \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_2\mathbf{C})][\mathbf{A}\mathbf{G}_1]\}^T + \{[\mathbf{A}(\mathbf{I} - \mathbf{G}_k\mathbf{C})] \\ &\quad \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_1\mathbf{C})]\}\mathbf{P}_0\{[\mathbf{A}(\mathbf{I} - \mathbf{G}_k\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_1\mathbf{C})]\}^T \end{aligned}$$

对于 $k \geq n+1$, 要证明 $\mathbf{P}_k > \mathbf{0}$, 只需说明由 $\mathbf{y}^T\mathbf{P}_k\mathbf{y} = 0$ 能得到 $\mathbf{y} = \mathbf{0}$ 就足够了。令 \mathbf{y} 是任意 n 维向量, 满足 $\mathbf{y}^T\mathbf{P}_k\mathbf{y} = 0$ 。由于 \mathbf{Q}, \mathbf{R} 和 \mathbf{P}_0 都是非负定的, 上式右边的每一项都必须为零。因此, 有

$$\mathbf{y}^T\boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T\mathbf{y} = 0 \quad (6.18)$$

$$\mathbf{y}^T[\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]\boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T[\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]^T\mathbf{y} = 0 \quad (6.19)$$

...

$$\begin{aligned} & \mathbf{y}^T \{ [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_2\mathbf{C})] \} \boldsymbol{\Gamma} \mathbf{Q} \boldsymbol{\Gamma}^T \\ & \cdot \{ [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_2\mathbf{C})] \}^T \mathbf{y} = 0 \end{aligned} \quad (6.20)$$

和

$$\mathbf{y}^T [\mathbf{A}\mathbf{G}_{k-1}\mathbf{C}] \mathbf{R} [\mathbf{A}\mathbf{G}_{k-1}]^T \mathbf{y} = 0 \quad (6.21)$$

...

$$\begin{aligned} & \mathbf{y}^T \{ [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_3\mathbf{C})] [\mathbf{A}\mathbf{G}_2] \} \mathbf{R} \\ & \cdot \{ [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_3\mathbf{C})] [\mathbf{A}\mathbf{G}_2] \}^T \mathbf{y} = 0 \end{aligned} \quad (6.22)$$

因为 $\mathbf{R} > \mathbf{0}$, 由式(6.21)和式(6.22), 有

$$\mathbf{y}^T \mathbf{A}\mathbf{G}_{k-1} = \mathbf{0} \quad (6.23)$$

...

$$\mathbf{y}^T [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_2\mathbf{C})] [\mathbf{A}\mathbf{G}_1] = \mathbf{0} \quad (6.24)$$

由 $\mathbf{Q} > \mathbf{0}$ 和式(6.18), 有

$$\mathbf{y}^T \boldsymbol{\Gamma} = \mathbf{0}$$

根据式(6.19)和式(6.23), 可得

$$\mathbf{y}^T \mathbf{A}\boldsymbol{\Gamma} = \mathbf{0}$$

以此类推, 可得当 $k \geq n+1$, 有

$$\mathbf{y}^T \mathbf{A}^j \boldsymbol{\Gamma} = \mathbf{0}, \quad j = 0, 1, \dots, n-1$$

即

$$\mathbf{y}^T \mathbf{M}_{A\boldsymbol{\Gamma}} \mathbf{y} = \mathbf{y}^T [\boldsymbol{\Gamma} \quad \mathbf{A}\boldsymbol{\Gamma} \quad \cdots \quad \mathbf{A}^{n-1}\boldsymbol{\Gamma}] \mathbf{y} = 0$$

因为系统是(完全)可控的, $\mathbf{M}_{A\boldsymbol{\Gamma}}$ 是满秩的, 必有 $\mathbf{y} = \mathbf{0}$ 。因此, 对所有 $k \geq n+1$, 有 $\mathbf{P}_k > \mathbf{0}$ 。这就完成了引理的证明。

现在, 反复应用式(6.9), 有

$$\mathbf{P} - \mathbf{P}_k = [\mathbf{A}(\mathbf{I} - \mathbf{GC})]^{k-n-1} (\mathbf{P} - \mathbf{P}_{n+1}) \mathbf{B}_k^T \quad (6.25)$$

式中

$$\mathbf{B}_k = [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \cdots [\mathbf{A}(\mathbf{I} - \mathbf{G}_{n+1}\mathbf{C})], \quad k = n+2, n+3, \dots$$

且有 $\mathbf{B}_{n+1} := \mathbf{I}$ 。为证明 $k \rightarrow \infty$ 时, $\mathbf{P}_k \rightarrow \mathbf{P}$, 只需证明 $k \rightarrow \infty$ 时, $[\mathbf{A}(\mathbf{I} - \mathbf{GC})]^{k-n-1} \rightarrow \mathbf{0}$, 且 \mathbf{B}_k 是有界的。为此目的, 我们需要下面的两个引理。

引理 6.7 设线性系统(6.1)可观测。则存在常数矩阵 \mathbf{M} , 满足

$$\mathbf{B}_k \mathbf{B}_k^T \leq \mathbf{M}, \quad k \geq n+1$$

因此, 如果 $\mathbf{B}_k = [b_{ij}^{(k)}]$, 则对于常数 m 和对所有的 i, j, k , 有

$$|b_{ij}^{(k)}| \leq m$$

由引理 6.1, 对于 $k \geq n+1$, $\mathbf{P}_k \leq \mathbf{W}$ 。反复应用引理 6.5, 有

$$\begin{aligned} \mathbf{W} & \geq \mathbf{P}_k \geq [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] \mathbf{P}_{k-1} [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]^T \\ & \geq [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})] [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-2}\mathbf{C})] \mathbf{P}_{k-2} [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-2}\mathbf{C})]^T [\mathbf{A}(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{C})]^T \\ & \cdots \\ & \geq \mathbf{B}_k \mathbf{P}_{n+1} \mathbf{B}_k^T \end{aligned}$$

因为 \mathbf{P}_{n+1} 是实对称正定矩阵,由引理 6.6,其所有特征值是实数,且实际上是正的。令 λ_{\min} 是 \mathbf{P}_{n+1} 的最小特征值,则有 $\mathbf{P}_{n+1} \geq \lambda_{\min} \mathbf{I}$ (见练习 6.3),从而有

$$\mathbf{W} \geq \mathbf{B}_k \lambda_{\min} \mathbf{I} \mathbf{B}_k^T = \lambda_{\min} \mathbf{B}_k \mathbf{B}_k^T$$

取 $\mathbf{M} = \lambda_{\min}^{-1} \mathbf{W}$,即可完成引理 6.7 的证明。

引理 6.8 设 λ 是 $\mathbf{A}(\mathbf{I} - \mathbf{G}\mathbf{C})$ 的任意特征值。如果系统(6.1)同时(完全)可控、可观,则 $|\lambda| < 1$ 。
■

注意到 λ 同时也是 $(\mathbf{I} - \mathbf{G}\mathbf{C})^T \mathbf{A}^T$ 的特征值。设 \mathbf{y} 是相应的特征向量,则

$$(\mathbf{I} - \mathbf{G}\mathbf{C})^T \mathbf{A}^T \mathbf{y} = \lambda \mathbf{y} \quad (6.26)$$

根据式(6.17),有

$$\bar{\mathbf{y}}^T \mathbf{P} \mathbf{y} = \bar{\lambda} \bar{\mathbf{y}}^T \mathbf{P} \lambda \mathbf{y} + \bar{\mathbf{y}}^T [\mathbf{AG}] \mathbf{R} [\mathbf{AG}]^T \mathbf{y} + \bar{\mathbf{y}}^T \boldsymbol{\Gamma} \mathbf{Q} \boldsymbol{\Gamma}^T \mathbf{y}$$

因此

$$(1 - |\lambda|^2) \bar{\mathbf{y}}^T \mathbf{P} \mathbf{y} = \bar{\mathbf{y}}^T [(\mathbf{AG}) \mathbf{R} (\mathbf{AG})^T + \boldsymbol{\Gamma} \mathbf{Q} \boldsymbol{\Gamma}^T] \mathbf{y}$$

因为右边是非负的,且 $\bar{\mathbf{y}}^T \mathbf{P} \mathbf{y} \geq 0$,必须有 $1 - |\lambda|^2 \geq 0$ 或 $|\lambda| \leq 1$ 。

假设 $|\lambda| = 1$,则

$$\bar{\mathbf{y}}^T [\mathbf{AG}] \mathbf{R} [\mathbf{AG}]^T \mathbf{y} = 0 \text{ 或 } \overline{[(\mathbf{AG})^T \mathbf{y}]}^T \mathbf{R} [(\mathbf{AG})^T \mathbf{y}] = 0$$

且

$$\bar{\mathbf{y}}^T \boldsymbol{\Gamma} \mathbf{Q} \boldsymbol{\Gamma}^T \mathbf{y} = 0 \text{ 或 } \overline{(\boldsymbol{\Gamma}^T \mathbf{y})}^T \mathbf{Q} (\boldsymbol{\Gamma}^T \mathbf{y}) = 0$$

因为 $\mathbf{Q} > \mathbf{0}$ 和 $\mathbf{R} > \mathbf{0}$,有

$$\mathbf{G}^T \mathbf{A}^T \mathbf{y} = \mathbf{0} \quad (6.27)$$

$$\boldsymbol{\Gamma}^T \mathbf{y} = \mathbf{0} \quad (6.28)$$

由式(6.26)可得 $\mathbf{A}^T \mathbf{y} = \lambda \mathbf{y}$ 。因此

$$\boldsymbol{\Gamma}^T (\mathbf{A}^j)^T \mathbf{y} = \lambda^j \boldsymbol{\Gamma}^T \mathbf{y} = \mathbf{0}, \quad j = 0, 1, \dots, n-1$$

可得

$$\mathbf{y}^T \mathbf{M}_{\mathbf{A}\boldsymbol{\Gamma}} = \mathbf{y}^T [\boldsymbol{\Gamma} \quad \mathbf{A}\boldsymbol{\Gamma} \quad \cdots \quad \mathbf{A}^{n-1} \boldsymbol{\Gamma}] = \mathbf{0}$$

分别取实部和虚部,有

$$[\operatorname{Re}(\mathbf{y})]^T \mathbf{M}_{\mathbf{A}\boldsymbol{\Gamma}} = \mathbf{0}, \quad [\operatorname{Im}(\mathbf{y})]^T \mathbf{M}_{\mathbf{A}\boldsymbol{\Gamma}} = \mathbf{0}$$

因为 $\mathbf{y} \neq \mathbf{0}$, $\operatorname{Re}(\mathbf{y})$ 和 $\operatorname{Im}(\mathbf{y})$ 至少有一个不为零。因此 $\mathbf{M}_{\mathbf{A}\boldsymbol{\Gamma}}$ 是行相关的,与完全可控的假设矛盾。因此 $|\lambda| < 1$ 。

这就完成了引理 6.8 的证明。

6.3 几何收敛

结合上面的结论,现在可以得到以下定理。

定理 6.1 设线性随机系统(6.1)同时(完全)可控可观。则对于任意初始状态 $x_0, P_0 := P_{0,-1} = \text{Var}(x_0)$ 非负定对称,当 $k \rightarrow \infty$ 时, $P_k := P_{k,k-1} \rightarrow P$, 其中对称矩阵 $P > 0$, 且独立于 x_0 。此外,收敛的阶是几何的,即

$$\text{tr}(P_k - P)(P_k - P)^T \leqslant Cr^k \quad (6.29)$$

式中, $0 < r < 1$ 和 $C > 0$ 独立于 k , 所以

$$\text{tr}(G_k - G)(G_k - G)^T \leqslant Cr^k \quad (6.30) \blacksquare$$

为证明该定理,令 $F = A(I - GC)$ 。根据引理 6.7 和式(6.25),存在非负定对称常值矩阵 Ω , 满足

$$\begin{aligned} (P_k - P)(P_k - P)^T &= F^{k-n-1}(P_{n+1} - P)B_kB_k^T(P_{n+1} - P)(F^{k-n-1})^T \\ &\leqslant F^{k-n-1}\Omega(F^{k-n-1})^T \end{aligned}$$

由引理 6.8, F 的所有特征值的绝对值都小于 1。因此 $k \rightarrow \infty$ 时, $F^k \rightarrow 0$, 以使得 $P_k \rightarrow P$ (见练习 6.4)。

另一方面,由引理 6.6, P 是正定对称的且独立于 P_0 。

根据引理 1.7 和引理 1.10, 有

$$\text{tr}(P_k - P)(P_k - P)^T \leqslant \text{tr}F^{k-n-1}(F^{k-n-1})^T \cdot \text{tr}\Omega \leqslant Cr^k$$

式中, $0 < r < 1$, 并且 r, c 与 k 无关。

为证明式(6.30),首先有

$$\begin{aligned} G_k - G &= P_k C^T (CP_k C^T + R)^{-1} - PC^T (CPC^T + R)^{-1} \\ &= (P_k - P)C^T (CP_k C^T + R)^{-1} \\ &\quad + PC^T [(CP_k C^T + R)^{-1} - (CPC^T + R)^{-1}] \\ &= (P_k - P)C^T (CP_k C^T + R)^{-1} + PC^T (CP_k C^T + R)^{-1} \\ &\quad \cdot [(CPC^T + R) - (CP_k C^T + R)] (CPC^T + R)^{-1} \\ &= (P_k - P)C^T (CP_k C^T + R)^{-1} \\ &\quad + PC^T (CP_k C^T + R)^{-1} C (P - P_k) C^T (CPC^T + R)^{-1} \end{aligned}$$

因为对任意 $n \times n$ 矩阵 A 和 B ,都有

$$(A + B)(A + B)^T \leqslant 2(AA^T + BB^T)$$

(见练习 6.5), 则

$$(G_k - G)(G_k - G)^T$$

$$\begin{aligned}
&\leq 2(\mathbf{P}_k - \mathbf{P})\mathbf{C}^T(\mathbf{C}\mathbf{P}_k\mathbf{C}^T + \mathbf{R})^{-1}(\mathbf{C}\mathbf{P}_k\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}(\mathbf{P}_k - \mathbf{P}) \\
&+ 2\mathbf{P}\mathbf{C}^T(\mathbf{C}\mathbf{P}_k\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}(\mathbf{P} - \mathbf{P}_k)\mathbf{C}^T(\mathbf{C}\mathbf{P}\mathbf{C}^T + \mathbf{R})^{-1} \\
&\cdot (\mathbf{C}\mathbf{P}\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}(\mathbf{P} - \mathbf{P}_k)\mathbf{C}^T(\mathbf{C}\mathbf{P}_k\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\mathbf{P}
\end{aligned} \tag{6.31}$$

因为 $\mathbf{P}_0 \leq \mathbf{P}_k$, 可得 $\mathbf{C}\mathbf{P}_0\mathbf{C}^T + \mathbf{R} \leq \mathbf{C}\mathbf{P}_k\mathbf{C}^T + \mathbf{R}$, 从而由引理 1.3 得

$$(\mathbf{C}\mathbf{P}_k\mathbf{C}^T + \mathbf{R})^{-1} \leq (\mathbf{C}\mathbf{P}_0\mathbf{C}^T + \mathbf{R})^{-1}$$

由引理 1.9, 得

$$\text{tr}(\mathbf{C}\mathbf{P}_k\mathbf{C}^T + \mathbf{R})^{-1}(\mathbf{C}\mathbf{P}_k\mathbf{C}^T + \mathbf{R})^{-1} \leq (\text{tr}(\mathbf{C}\mathbf{P}_0\mathbf{C}^T + \mathbf{R})^{-1})^2$$

最后, 根据引理 1.7, 由式(6.31)可得

$$\begin{aligned}
\text{tr}(\mathbf{G}_k - \mathbf{G})(\mathbf{G}_k - \mathbf{G})^T &\leq 2\text{tr}(\mathbf{P}_k - \mathbf{P})(\mathbf{P}_k - \mathbf{P})^T \\
&\quad \cdot \text{tr}\mathbf{C}^T\mathbf{C}(\text{tr}(\mathbf{C}\mathbf{P}_0\mathbf{C}^T + \mathbf{R})^{-1})^2 + 2\text{tr}\mathbf{P}\mathbf{P}^T \\
&\quad \cdot \text{tr}\mathbf{C}^T\mathbf{C}(\text{tr}(\mathbf{C}\mathbf{P}_0\mathbf{C}^T + \mathbf{R})^{-1})^2 \cdot \text{tr}\mathbf{C}\mathbf{C}^T \cdot \text{tr}(\mathbf{P} - \mathbf{P}_k)(\mathbf{P} - \mathbf{P}_k)^T \\
&\quad \cdot \text{tr}\mathbf{C}^T\mathbf{C} \cdot \text{tr}(\mathbf{C}\mathbf{P}\mathbf{C}^T + \mathbf{R})^{-1}(\mathbf{C}\mathbf{P}\mathbf{C}^T + \mathbf{R})^{-1} \\
&\leq C_1 \text{tr}(\mathbf{P}_k - \mathbf{P})(\mathbf{P}_k - \mathbf{P})^T \\
&\leq Cr^k
\end{aligned}$$

式中, C_1 和 C 是与 k 无关的常值。

这就完成了定理 6.1 的证明。

下面的结果说明 \vec{x}_k 是 \mathbf{x}_k 的渐近最优估计。

定理 6.2 设线性随机系统(6.1)同时(完全)可控可观, 那么

$$\lim_{k \rightarrow \infty} \| \mathbf{x}_k - \vec{x}_k \|_n^2 = (\mathbf{P}^{-1} + \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C})^{-1} = \lim_{k \rightarrow \infty} \| \mathbf{x}_k - \hat{x}_k \|_n^2 \quad \blacksquare$$

定理 6.2 的第二个等式可以很容易证明。利用引理 1.2(矩阵求逆引理), 上式可写为

$$\begin{aligned}
\lim_{k \rightarrow \infty} \| \mathbf{x}_k - \hat{x}_k \|_n^2 &= \lim_{k \rightarrow \infty} \mathbf{P}_{k,k} \\
&= \lim_{k \rightarrow \infty} (\mathbf{I} - \mathbf{G}_k \mathbf{C}) \mathbf{P}_{k,k-1} \\
&= (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P} \\
&= \mathbf{P} - \mathbf{P}\mathbf{C}^T(\mathbf{C}\mathbf{P}\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\mathbf{P} \\
&= (\mathbf{P}^{-1} + \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C})^{-1} > 0
\end{aligned}$$

而证明第一个等式, 等价于证明 $k \rightarrow \infty$ 时, $\| \mathbf{x}_k - \vec{x}_k \|_n^2 \rightarrow (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P}$ 。首先有

$$\begin{aligned}
\mathbf{x}_k - \vec{x}_k &= (\mathbf{A}\mathbf{x}_{k-1} + \Gamma \underline{\xi}_{k-1}) - (\mathbf{A}\vec{x}_{k-1} + \mathbf{G}\mathbf{v}_k - \mathbf{G}\mathbf{C}\mathbf{A}\vec{x}_{k-1}) \\
&= (\mathbf{A}\mathbf{x}_{k-1} + \Gamma \underline{\xi}_{k-1}) - \mathbf{A}\vec{x}_{k-1} - \mathbf{G}(\mathbf{C}\mathbf{A}\mathbf{x}_{k-1} + \mathbf{C}\Gamma \underline{\xi}_{k-1} + \underline{\eta}_k) + \mathbf{G}\mathbf{C}\mathbf{A}\vec{x}_{k-1} \\
&= (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A}(\mathbf{x}_{k-1} - \vec{x}_{k-1}) + (\mathbf{I} - \mathbf{G}\mathbf{C})\Gamma \underline{\xi}_{k-1} - \mathbf{G}\underline{\eta}_k
\end{aligned} \tag{6.32}$$

又

$$\langle \underline{x}_{k-1} - \vec{x}_{k-1}, \underline{\xi}_{k-1} \rangle = 0 \quad (6.33)$$

$$\langle \underline{x}_{k-1} - \vec{x}_{k-1}, \underline{\eta}_k \rangle = 0 \quad (6.34)$$

(见练习 6.6)。有

$$\begin{aligned} \| \underline{x}_k - \vec{x}_k \|_n^2 &= (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A} \| \underline{x}_{k-1} - \vec{x}_{k-1} \|_n^2 \mathbf{A}^\top (\mathbf{I} - \mathbf{G}\mathbf{C})^\top \\ &\quad + (\mathbf{I} - \mathbf{G}\mathbf{C}) \Gamma Q \Gamma^\top (\mathbf{I} - \mathbf{G}\mathbf{C})^\top + \mathbf{G}\mathbf{R}\mathbf{G}^\top \end{aligned} \quad (6.35)$$

另一方面,可以证明

$$\mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k \mathbf{C}) \mathbf{A} \mathbf{P}_{k-1,k-1} \mathbf{A}^\top (\mathbf{I} - \mathbf{G}_k \mathbf{C})^\top + (\mathbf{I} - \mathbf{G}_k \mathbf{C}) \Gamma Q \Gamma^\top (\mathbf{I} - \mathbf{G}_k \mathbf{C})^\top + \mathbf{G}_k \mathbf{R} \mathbf{G}_k^\top \quad (6.36)$$

(见练习 6.7)。由于当 $k \rightarrow \infty$ 时, $\mathbf{P}_{k,k} = (\mathbf{I} - \mathbf{G}_k \mathbf{C}) \mathbf{P}_{k,k-1} \rightarrow (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P}$, 取极限得到

$$\begin{aligned} (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P} &= (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A} [(\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P}] \mathbf{A}^\top (\mathbf{I} - \mathbf{G}\mathbf{C})^\top \\ &\quad + (\mathbf{I} - \mathbf{G}\mathbf{C}) \Gamma Q \Gamma^\top (\mathbf{I} - \mathbf{G}\mathbf{C})^\top + \mathbf{G}\mathbf{R}\mathbf{G}^\top \end{aligned} \quad (6.37)$$

由式(6.35)减去(6.37),得

$$\begin{aligned} &\| \underline{x}_k - \vec{x}_k \|_n^2 - (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P} \\ &= (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A} [\| \underline{x}_{k-1} - \vec{x}_{k-1} \|_n^2 - (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P}] \mathbf{A}^\top (\mathbf{I} - \mathbf{G}\mathbf{C})^\top \end{aligned}$$

反复应用该公式 $k-1$ 次,得

$$\begin{aligned} &\| \underline{x}_k - \vec{x}_k \|_n^2 - (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P} \\ &= [(\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A}]^k [\| \underline{x}_0 - \vec{x}_0 \|_n^2 - (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P}] [\mathbf{A}^\top (\mathbf{I} - \mathbf{G}\mathbf{C})^\top]^k \end{aligned}$$

最后,模仿引理 6.8 的证明,可以得到 $(\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A}$ 的所有特征值的绝对值都小于 1 (见练习 6.8)。因此,根据练习 6.4,当 $k \rightarrow \infty$ 时,有 $\| \underline{x}_k - \vec{x}_k \|_n^2 - (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{P} \rightarrow 0$ 。这就完成了定理的证明。

下面说明误差 $\hat{x}_k - \vec{x}_k$ 也以指数速率趋于零。

定理 6.3 设线性随机系统(6.1)同时(完全)可控可观,则存在分别独立于 k 的实数 r ($0 < r < 1$) 和正的常数 C ,使得

$$\text{tr } \| \hat{x}_k - \vec{x}_k \|_n^2 \leqslant Cr^k$$

证明如下。

记 $\underline{\epsilon}_k := \hat{x}_k - \vec{x}_k$, $\underline{\delta}_k := x_k - \hat{x}_k$ 。由下列两个恒等式

$$\begin{aligned} \hat{x}_k &= \mathbf{A} \hat{x}_{k-1} + \mathbf{G}_k (\underline{v}_k - \mathbf{C} \mathbf{A} \hat{x}_{k-1}) \\ &= \mathbf{A} \hat{x}_{k-1} + \mathbf{G} (\underline{v}_k - \mathbf{C} \mathbf{A} \hat{x}_{k-1}) + (\mathbf{G}_k - \mathbf{G})(\underline{v}_k - \mathbf{C} \mathbf{A} \hat{x}_{k-1}) \\ \vec{x}_k &= \mathbf{A} \vec{x}_{k-1} + \mathbf{G} (\underline{v}_k - \mathbf{C} \mathbf{A} \vec{x}_{k-1}) \end{aligned}$$

得

$$\underline{\epsilon}_k = \hat{x}_k - \vec{x}_k$$

$$\begin{aligned}
&= \mathbf{A}(\hat{\mathbf{x}}_{k-1} - \vec{\mathbf{x}}_{k-1}) - \mathbf{G}\mathbf{C}\mathbf{A}(\hat{\mathbf{x}}_{k-1} - \vec{\mathbf{x}}_{k-1}) \\
&\quad + (\mathbf{G}_k - \mathbf{G})(\mathbf{C}\mathbf{A}\vec{\mathbf{x}}_{k-1} + \mathbf{C}\Gamma\hat{\boldsymbol{\xi}}_{k-1} + \underline{\boldsymbol{\eta}}_k - \mathbf{C}\mathbf{A}\hat{\mathbf{x}}_{k-1}) \\
&= (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A}\underline{\boldsymbol{\epsilon}}_{k-1} + (\mathbf{G}_k - \mathbf{G})(\mathbf{C}\mathbf{A}\underline{\boldsymbol{\delta}}_{k-1} + \mathbf{C}\Gamma\hat{\boldsymbol{\xi}}_{k-1} + \underline{\boldsymbol{\eta}}_k)
\end{aligned}$$

又

$$\left\{
\begin{array}{l}
\langle \underline{\boldsymbol{\epsilon}}_{k-1}, \hat{\boldsymbol{\xi}}_{k-1} \rangle = \mathbf{0}, \langle \underline{\boldsymbol{\epsilon}}_{k-1}, \underline{\boldsymbol{\eta}}_k \rangle = \mathbf{0} \\
\langle \underline{\boldsymbol{\delta}}_{k-1}, \hat{\boldsymbol{\xi}}_{k-1} \rangle = \mathbf{0}, \langle \underline{\boldsymbol{\delta}}_{k-1}, \underline{\boldsymbol{\eta}}_k \rangle = \mathbf{0}
\end{array}
\right. \quad (6.38)$$

及 $\langle \hat{\boldsymbol{\xi}}_{k-1}, \underline{\boldsymbol{\eta}}_k \rangle = \mathbf{0}$ (见练习 6.9), 可得

$$\begin{aligned}
\| \underline{\boldsymbol{\epsilon}}_k \|_n^2 &= [(\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A}] \| \underline{\boldsymbol{\epsilon}}_{k-1} \|_n^2 [(\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A}]^\top \\
&\quad + (\mathbf{G}_k - \mathbf{G})\mathbf{C}\mathbf{A} \| \underline{\boldsymbol{\delta}}_{k-1} \|_n^2 \mathbf{A}^\top \mathbf{C}^\top (\mathbf{G}_k - \mathbf{G})^\top \\
&\quad + (\mathbf{G}_k - \mathbf{G})\mathbf{C}\mathbf{T}\mathbf{Q}\mathbf{F}^\top \mathbf{C}^\top (\mathbf{G}_k - \mathbf{G})^\top + (\mathbf{G}_k - \mathbf{G})\mathbf{R}(\mathbf{G}_k - \mathbf{G})^\top \\
&\quad + (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A} \langle \underline{\boldsymbol{\epsilon}}_{k-1}, \underline{\boldsymbol{\delta}}_{k-1} \rangle \mathbf{A}^\top \mathbf{C}^\top (\mathbf{G}_k - \mathbf{G})^\top \\
&\quad + (\mathbf{G}_k - \mathbf{G})\mathbf{C}\mathbf{A} \langle \underline{\boldsymbol{\delta}}_{k-1}, \underline{\boldsymbol{\epsilon}}_{k-1} \rangle \mathbf{A}^\top (\mathbf{I} - \mathbf{G}\mathbf{C})^\top \\
&= \mathbf{F} \| \underline{\boldsymbol{\epsilon}}_{k-1} \|_n^2 \mathbf{F}^\top + (\mathbf{G}_k - \mathbf{G}) \underline{\boldsymbol{\Omega}}_{k-1} (\mathbf{G}_k - \mathbf{G})^\top \\
&\quad + \mathbf{F}\mathbf{B}_{k-1} (\mathbf{G}_k - \mathbf{G})^\top + (\mathbf{G}_k - \mathbf{G})\mathbf{B}_{k-1}^\top \mathbf{F}^\top \quad (6.39)
\end{aligned}$$

式中

$$\begin{aligned}
\mathbf{F} &= (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A} \\
\underline{\boldsymbol{\Omega}}_{k-1} &= \langle \underline{\boldsymbol{\epsilon}}_{k-1}, \underline{\boldsymbol{\delta}}_{k-1} \rangle \mathbf{A}^\top \mathbf{C}^\top \\
\mathbf{B}_{k-1} &= \mathbf{C}\mathbf{A} \| \underline{\boldsymbol{\delta}}_{k-1} \|_n^2 \mathbf{A}^\top \mathbf{C}^\top + \mathbf{C}\mathbf{T}\mathbf{Q}\mathbf{F}^\top \mathbf{C}^\top + \mathbf{R}
\end{aligned}$$

反复应用式(6.39), 可得

$$\begin{aligned}
\| \underline{\boldsymbol{\epsilon}}_k \|_n^2 &= \mathbf{F}^k \| \underline{\boldsymbol{\epsilon}}_0 \|_n^2 (\mathbf{F}^k)^\top + \sum_{i=0}^{k-1} \mathbf{F}^i (\mathbf{G}_{k-i} - \mathbf{G}) \underline{\boldsymbol{\Omega}}_{k-1-i} (\mathbf{G}_{k-i} - \mathbf{G})^\top (\mathbf{F}^i)^\top \\
&\quad + \sum_{i=0}^{k-1} \mathbf{F}^i [\mathbf{F}\mathbf{B}_{k-1-i} (\mathbf{G}_{k-i} - \mathbf{G})^\top + (\mathbf{G}_{k-i} - \mathbf{G})\mathbf{B}_{k-1-i}^\top \mathbf{F}^\top] (\mathbf{F}^i)^\top \quad (6.40)
\end{aligned}$$

另一方面, 由于 \mathbf{B}_i 的各分量一致有界(见练习 6.10), 可以证明, 通过应用引理 1.6、引理 1.7 和引理 1.10, 及定理 6.1, 对某些独立于 k, i 的 $r_1, 0 < r_1 < 1$, 及正常数 C_1 , 有

$$\mathbf{tr}[\mathbf{F}\mathbf{B}_{k-1-i} (\mathbf{G}_{k-i} - \mathbf{G})^\top + (\mathbf{G}_{k-i} - \mathbf{G})\mathbf{B}_{k-1-i}^\top \mathbf{F}^\top] \leqslant C_1 r_1^{k-i+1} \quad (6.41)$$

(见练习 6.11)。

再次应用引理 1.7 和引理 1.10 以及定理 6.1, 可得

$$\begin{aligned}
\text{tr} \parallel \underline{\epsilon}_k \parallel_n^2 &\leqslant \text{tr} \parallel \underline{\epsilon}_0 \parallel_n^2 + \text{tr} \mathbf{F}^k (\mathbf{F}^k)^T + \sum_{i=0}^{k-1} \text{tr} \mathbf{F}^i (\mathbf{F}^i)^T \\
&\quad \cdot \text{tr} (\mathbf{G}_{k-i} - \mathbf{G}) (\mathbf{G}_{k-i} - \mathbf{G})^T \cdot \text{tr} \underline{\Omega}_{k-1-i} \\
&+ \sum_{i=0}^{k-1} \text{tr} \mathbf{F}^i (\mathbf{F}^i)^T \cdot \text{tr} [\mathbf{F} \mathbf{B}_{k-1-i} (\mathbf{G}_{k-i} - \mathbf{G})^T + (\mathbf{G}_{k-i} - \mathbf{G}) \mathbf{B}_{k-1-i}^T \mathbf{F}^T] \\
&\leqslant \text{tr} \parallel \underline{\epsilon}_0 \parallel_n^2 C_2 r_2^k + \sum_{i=0}^{k-1} C_3 r_3^i C_4 r_4^{k-i} + \sum_{i=0}^{k-1} C_5 r_5^i C_1 r_1^{k-i+1} \\
&\leqslant p(k) r_6^k
\end{aligned} \tag{6.42}$$

式中, $0 < r_2, r_3, r_4, r_5 < 1$, $r_6 = \max(r_1, r_2, r_3, r_4, r_5) < 1$, 以及 C_2, C_3, C_4, C_5 是分别独立于 k 和 i 的正常值, 且 $p(k)$ 是 k 的多项式。因此, 存在一个独立于 k 的实数 $r, r_6 < r < 1$, 和一个正常数 C , 满足 $p(k)(r_6/r)^k \leqslant C$, 使得

$$\text{tr} \parallel \underline{\epsilon}_k \parallel_n^2 \leqslant C r^k$$

这就完成了定理的证明。

6.4 实时应用

再次研究跟踪模型(3.26), 其状态空间模型:

$$\begin{cases} \mathbf{x}_{k+1} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \underline{\xi}_k \\ \mathbf{v}_k = [1 \ 0 \ 0] \mathbf{x}_k + \eta_k \end{cases} \tag{6.43}$$

式中 $h > 0$ 为采样时间, $\{\underline{\xi}_k\}$ 和 $\{\eta_k\}$ 是零均值高斯白噪声序列, 且满足下列假设

$$E(\underline{\xi}_k \underline{\xi}_l^T) = \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix} \delta_{kl}, \quad E(\eta_k \eta_l) = \sigma_m \delta_{kl}$$

$$E(\underline{\xi}_k \eta_l) = \mathbf{0}, \quad E(\underline{\xi}_k \mathbf{x}_0^T) = \mathbf{0}, \quad E(\mathbf{x}_0 \eta_k) = \mathbf{0}$$

且 $\sigma_p, \sigma_v, \sigma_a \geqslant 0$, $\sigma_p + \sigma_v + \sigma_a > 0$ 和 $\sigma_m > 0$ 。

因为矩阵

$$\begin{aligned} \mathbf{M}_{A\Gamma} &= [\Gamma \ A\Gamma \ A^2\Gamma] \\ &= \begin{bmatrix} 1 & 0 & 0 & 1 & h & h^2/2 & 1 & 2h & 2h^2 \\ 0 & 1 & 0 & 0 & 1 & h & 0 & 1 & 2h \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

$$\mathbf{N}_{CA} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & h & h^2/2 \\ 1 & 2h & 2h^2 \end{bmatrix}$$

都满秩,所以系统(6.43)是完全可控可观的。根据定理6.1,存在正定对称矩阵 \mathbf{P} ,使得

$$\lim_{k \rightarrow \infty} \mathbf{P}_{k+1,k} = \mathbf{P}$$

式中

$$\mathbf{P}_{k+1,k} = \mathbf{A}(\mathbf{I} - \mathbf{G}_k \mathbf{C}^T) \mathbf{P}_{k,k-1} \mathbf{A}^T + \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix}$$

$$\mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C} (\mathbf{C}^T \mathbf{P}_{k,k-1} \mathbf{C} + \sigma_m)^{-1}$$

将 \mathbf{G}_k 代入上面 $\mathbf{P}_{k+1,k}$ 的表达式,然后取极限,可得下面的矩阵Riccati方程

$$\mathbf{P} = \mathbf{A}[\mathbf{P} - \mathbf{P}\mathbf{C}(\mathbf{C}^T\mathbf{P}\mathbf{C} + \sigma_m)^{-1}\mathbf{C}^T\mathbf{P}] \mathbf{A}^T + \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix} \quad (6.44)$$

现在求解这个正定矩阵 \mathbf{P} 的矩阵Riccati方程,得到极限卡尔曼增益:

$$\mathbf{G} = \mathbf{P}\mathbf{C}/(\mathbf{C}^T\mathbf{P}\mathbf{C} + \sigma_m)$$

及极限(稳态)卡尔曼滤波方程:

$$\begin{cases} \vec{x}_{k+1} = \mathbf{A}\vec{x}_k + \mathbf{G}(\mathbf{v}_k - \mathbf{C}\mathbf{A}\vec{x}_k) \\ \vec{x}_0 = E(\mathbf{x}_0) \end{cases} \quad (6.45)$$

因为矩阵Riccati方程(6.44)可以在滤波过程进行之前进行求解,该极限卡尔曼滤波给出了一个非常有效的实时跟踪器。由定理6.3可知,估计 \vec{x}_k 和最优估计 \hat{x}_k 非常接近。

练习

6.1 证明在 $E(\tilde{\mathbf{x}}_{k-1})=E(\mathbf{x}_{k-1})$ 的条件下,式(6.7)的估计 $\tilde{\mathbf{x}}_{k-1}$ 是 \mathbf{x}_{k-1} 的无偏估计。

6.2 证明: $\frac{d}{ds}\mathbf{A}^{-1}(s) = -\mathbf{A}^{-1}(s)\left[\frac{d}{ds}\mathbf{A}(s)\right]\mathbf{A}^{-1}(s)$ 。

6.3 证明若 λ_{\min} 是 \mathbf{P} 的最小特征值,则 $\mathbf{P} \geq \lambda_{\min} \mathbf{I}$ 。类似地,若 λ_{\max} 是 \mathbf{P} 的最大特征值,则 $\mathbf{P} \leq \lambda_{\max} \mathbf{I}$ 。

6.4 设 \mathbf{F} 是一个 $n \times n$ 阶矩阵, 并假设 \mathbf{F} 的所有特征值的绝对值都小于 1。证明当 $k \rightarrow \infty$ 时, $\mathbf{F}^k \rightarrow \mathbf{0}$ 。

6.5 证明对于任意 $n \times n$ 阶矩阵 \mathbf{A} 和 \mathbf{B} , 有 $(\mathbf{A} + \mathbf{B})(\mathbf{A} + \mathbf{B})^T \leq 2(\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T)$ 。

6.6 设 $\{\xi_k\}$ 和 $\{\eta_k\}$ 分别是系统和量测的零均值高斯白噪声序列, \vec{x}_k 由式(6.4)定义。证明:

$$\langle \mathbf{x}_{k-1} - \vec{x}_{k-1}, \xi_{k-1} \rangle = \mathbf{0}$$

$$\langle \mathbf{x}_{k-1} - \vec{x}_{k-1}, \eta_k \rangle = \mathbf{0}$$

6.7 证明对于卡尔曼增益 \mathbf{G}_k , 有 $-(\mathbf{I} - \mathbf{G}_k \mathbf{C}) \mathbf{P}_{k,k-1} \mathbf{C}^T \mathbf{G}_k^T + \mathbf{G}_k \mathbf{R} \mathbf{G}_k^T = \mathbf{0}$, 并应用此公式, 证明:

$$\begin{aligned} \mathbf{P}_{k,k} &= (\mathbf{I} - \mathbf{G}_k \mathbf{C}) \mathbf{A} \mathbf{P}_{k-1,k-1} \mathbf{A}^T (\mathbf{I} - \mathbf{G}_k \mathbf{C})^T \\ &\quad + (\mathbf{I} - \mathbf{G}_k \mathbf{C}) \mathbf{P} \mathbf{Q}_k \mathbf{P}^T (\mathbf{I} - \mathbf{G}_k \mathbf{C})^T + \mathbf{G}_k \mathbf{R} \mathbf{G}_k^T. \end{aligned}$$

6.8 模仿引理 6.8 的证明过程, 证明 $(\mathbf{I} - \mathbf{G} \mathbf{C}) \mathbf{A}$ 的所有特征值的绝对值都小于 1。

6.9 设 $\underline{\epsilon}_k = \hat{x}_k - \vec{x}_k$, 其中 \vec{x}_k 由式(6.4)定义, 令 $\underline{\delta}_k = \mathbf{x}_k - \vec{x}_k$, 证明

$$\langle \underline{\epsilon}_{k-1}, \xi_{k-1} \rangle = \mathbf{0}, \quad \langle \underline{\epsilon}_{k-1}, \eta_k \rangle = \mathbf{0}$$

$$\langle \underline{\delta}_{k-1}, \xi_{k-1} \rangle = \mathbf{0}, \quad \langle \underline{\delta}_{k-1}, \eta_k \rangle = \mathbf{0}$$

式中 $\{\xi_k\}$ 和 $\{\eta_k\}$ 分别是系统和量测的零均值高斯白噪声序列。

6.10 设

$$\mathbf{B}_j = \langle \underline{\epsilon}_j, \underline{\delta}_j \rangle \mathbf{A}^T \mathbf{C}^T, \quad j = 0, 1, \dots$$

式中, $\underline{\epsilon}_j = \hat{x}_j - \vec{x}_j$, $\underline{\delta}_j = \mathbf{x}_j - \hat{x}_j$, \vec{x}_j 由式(6.4)定义。证明 \mathbf{B}_j 的各分量一致有界。

6.11 推导式(6.41)。

6.12 给出下面标量系统的极限(稳态)卡尔曼滤波的具体算法:

$$\begin{cases} x_{k+1} = ax_k + \gamma \xi_k \\ v_k = cx_k + \eta_k \end{cases}$$

式中 a, γ 和 c 是常数, $\{\xi_k\}$ 、 $\{\eta_k\}$ 分别是方差为 q 和 r 的零均值高斯白噪声序列。

第 7 章

序贯算法和平方根算法

很明显,卡尔曼滤波过程中唯一耗时的操作就是卡尔曼增益矩阵的计算:

$$\mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1}$$

而实时性是卡尔曼滤波需要考虑的首要问题。因此在滤波过程中,极其重要的是在每个时刻不必直接对矩阵求逆就能计算出 \mathbf{G}_k ,或者高效精确地执行修正操作,无论该操作是否包含矩阵求逆过程。我们首先讨论的序贯算法避免了对矩阵 $(\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)$ 的直接求逆;其次研究的平方根算法只对三角矩阵进行求逆,并且求解可能非常大或者非常小数字的平方根时,可以提高计算精度。最后将这两种算法组合,得到一种可以满足实时应用的高效计算方案。

7.1 序贯算法

如果正定阵 \mathbf{R}_k 是对角阵,即

$$\mathbf{R}_k = \text{diag}[r_k^1, \dots, r_k^q]$$

式中 $r_k^1, \dots, r_k^q > 0$,那么序贯算法的效率很高。如果 \mathbf{R}_k 不是对角阵,那么可以确定正交阵 \mathbf{T}_k ,使得转换矩阵 $\mathbf{T}_k^T \mathbf{R}_k \mathbf{T}_k$ 是对角阵。

这样,状态空间描述的观测方程:

$$\mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \underline{\eta}_k$$

可以转换为

$$\tilde{\mathbf{v}}_k = \tilde{\mathbf{C}}_k \mathbf{x}_k + \tilde{\underline{\eta}}_k$$

式中 $\tilde{\mathbf{v}}_k = \mathbf{T}_k^T \mathbf{v}_k$, $\tilde{\mathbf{C}}_k = \mathbf{T}_k^T \mathbf{C}_k$, $\tilde{\underline{\eta}}_k = \mathbf{T}_k^T \underline{\eta}_k$, 并且

$$\text{Var}(\tilde{\eta}_k) = \mathbf{T}_k^T \mathbf{R}_k \mathbf{T}_k$$

在接下来的讨论中,假设 \mathbf{R}_k 是对角阵。我们所感兴趣的只是计算卡尔曼增益矩阵 \mathbf{G}_k 和 k 时刻状态向量 \mathbf{x}_k 的最优估计 $\hat{\mathbf{x}}_{k|k}$ 。只要不引起混淆,可以去掉下标 k ,例如:

$$\mathbf{v}_k = \begin{bmatrix} \mathbf{v}^1 \\ \vdots \\ \mathbf{v}^q \end{bmatrix}_{q \times 1}, \quad \mathbf{C}_k^T = [\mathbf{c}^1 \quad \cdots \quad \mathbf{c}^q]_{n \times q}$$

$$\mathbf{R}_k = \text{diag}[r^1, \dots, r^q]$$

序贯算法描述如下。

定理 7.1 令 k 是固定的,设

$$\mathbf{P}^0 = \mathbf{P}_{k,k-1}, \quad \hat{\mathbf{x}}^0 = \hat{\mathbf{x}}_{k|k-1} \quad (7.1)$$

且对于 $i=1, \dots, q$,计算

$$\begin{cases} \mathbf{g}^i = \frac{1}{(\mathbf{c}^i)^T \mathbf{P}^{i-1} \mathbf{c}^i + r^i} \mathbf{P}^{i-1} \mathbf{c}^i \\ \hat{\mathbf{x}}^i = \hat{\mathbf{x}}^{i-1} + [\mathbf{v}^i - (\mathbf{c}^i)^T \hat{\mathbf{x}}^{i-1}] \mathbf{g}^i \\ \mathbf{P}^i = \mathbf{P}^{i-1} - \mathbf{g}^i (\mathbf{c}^i)^T \mathbf{P}^{i-1} \end{cases} \quad (7.2)$$

由此可得

$$\mathbf{G}_k = \mathbf{P}^q \mathbf{C}_k^T \mathbf{R}_k^{-1} \quad (7.3)$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}^q \quad (7.4)$$

(见图 7.1)。 ■

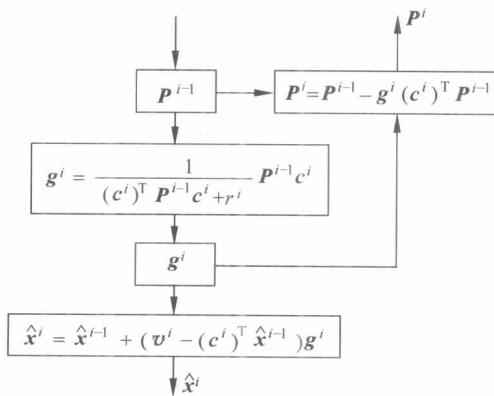


图 7.1 序贯算法

定理证明如下。

为了证明式(7.3),首先需证明:

$$\mathbf{G}_k = \mathbf{P}_{k,k} \mathbf{C}_k^T \mathbf{R}_k^{-1} \quad (7.5)$$

该式可由滤波方程得到

$$\begin{aligned}\mathbf{P}_{k,k} &= (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \\ &= \mathbf{P}_{k,k-1} - \mathbf{G}_k \mathbf{C}_k \mathbf{P}_{k,k-1}\end{aligned}$$

故

$$\begin{aligned}\mathbf{G}_k &= \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ &= (\mathbf{P}_{k,k} + \mathbf{G}_k \mathbf{C}_k \mathbf{P}_{k,k-1}) \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1}\end{aligned}$$

或者

$$\mathbf{G}_k (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k) = (\mathbf{P}_{k,k} + \mathbf{G}_k \mathbf{C}_k \mathbf{P}_{k,k-1}) \mathbf{C}_k^T$$

由该式可求得(7.5)。

因此,如果证得式(7.6)成立,即可证得式(7.3)成立。再来证

$$\mathbf{P}_{k,k} = \mathbf{P}^q \quad (7.6)$$

该等式的直接证明看来是不可得的,需借助矩阵求逆引理(引理1.2)。令 $\epsilon > 0$, 并设置正定阵 $\mathbf{P}_\epsilon^0 = \mathbf{P}_{k,k-1} + \epsilon \mathbf{I}$ 。同理,设置

$$\begin{aligned}\mathbf{P}_\epsilon^i &= \mathbf{P}_\epsilon^{i-1} - \mathbf{g}_\epsilon^i (\mathbf{c}^i)^T \mathbf{P}_\epsilon^{i-1} \\ \mathbf{g}_\epsilon^i &= \frac{1}{(\mathbf{c}^i)^T \mathbf{P}_\epsilon^{i-1} \mathbf{c}^i + r^i} \mathbf{P}_\epsilon^{i-1} \mathbf{c}^i\end{aligned}$$

由 $i=1$ 开始用归纳论证和矩阵求逆引理式(1.3)。可以看出,矩阵

$$\mathbf{P}_\epsilon^i = \mathbf{P}_\epsilon^{i-1} - \mathbf{P}_\epsilon^{i-1} \mathbf{c}^i [(\mathbf{c}^i)^T \mathbf{P}_\epsilon^{i-1} \mathbf{c}^i + r^i]^{-1} (\mathbf{c}^i)^T \mathbf{P}_\epsilon^{i-1}$$

是可逆的,且有

$$(\mathbf{P}_\epsilon^i)^{-1} = (\mathbf{P}_\epsilon^{i-1})^{-1} + \mathbf{c}^i (r^i)^{-1} (\mathbf{c}^i)^T$$

对 $i=q, q-1, \dots, 1$,依次应用所有这些方程,可得

$$\begin{aligned}(\mathbf{P}_\epsilon^q)^{-1} &= (\mathbf{P}_\epsilon^{q-1})^{-1} + \mathbf{c}^q (r^q)^{-1} (\mathbf{c}^q)^T \\ &= \dots \\ &= (\mathbf{P}_\epsilon^0)^{-1} + \sum_{i=1}^q \mathbf{c}^i (r^i)^{-1} (\mathbf{c}^i)^T \\ &= (\mathbf{P}_\epsilon^0)^{-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k\end{aligned}$$

另一方面,再次使用矩阵求逆引理,可证矩阵

$$\tilde{\mathbf{P}}_\epsilon := \mathbf{P}_\epsilon^0 - \mathbf{P}_\epsilon^0 \mathbf{C}_k (\mathbf{C}_k \mathbf{P}_\epsilon^0 \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \mathbf{P}_\epsilon^0$$

也是可逆的,且

$$\tilde{\mathbf{P}}_\epsilon^{-1} = (\mathbf{P}_\epsilon^0)^{-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k$$

因此,有 $\tilde{\mathbf{P}}_\epsilon^{-1} = (\mathbf{P}_\epsilon^q)^{-1}$, 所以

$$\tilde{\mathbf{P}}_{\epsilon} = \mathbf{P}_{\epsilon}^q$$

随着 $\epsilon \rightarrow 0$, 由卡尔曼滤波方程, 可得

$$\begin{aligned}\tilde{\mathbf{P}}_{\epsilon} &\rightarrow \mathbf{P}^0 - \mathbf{P}^0 \mathbf{C}_k (\mathbf{C}_k \mathbf{P}^0 \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \mathbf{P}^0 \\ &= (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} = \mathbf{P}_{k,k}\end{aligned}$$

从这个定义中, 可以得到, $\epsilon \rightarrow 0$ 时,

$$\mathbf{P}_{\epsilon}^q \rightarrow \mathbf{P}^q$$

这意味着式(7.6)成立, 从而式(7.3)得证。

为证明式(7.4), 首先注意到

$$\mathbf{g}^i = \mathbf{P}^i \mathbf{c}^i (r^i)^{-1} \quad (7.7)$$

由式(7.2)中第三个方程, 可得

$$\mathbf{P}^{i-1} = \mathbf{P}^i + \mathbf{g}^i (\mathbf{c}^i)^T \mathbf{P}^{i-1}$$

由式(7.2)中的第一个方程, 可得

$$\mathbf{g}^i = \frac{1}{(\mathbf{c}^i)^T \mathbf{P}^{i-1} \mathbf{c}^i + r^i} (\mathbf{P}^i + \mathbf{g}^i (\mathbf{c}^i)^T \mathbf{P}^{i-1}) \mathbf{c}^i$$

其简化形式是式(7.7)。

对于任意的 $i, 0 \leq i \leq q-1$, 由式(7.2)中的第三个方程, 可得

$$\begin{aligned}\mathbf{P}^q &= (\mathbf{I} - \mathbf{g}^q (\mathbf{c}^q)^T) \mathbf{P}^{q-1} \\ &= \dots \dots \\ &= (\mathbf{I} - \mathbf{g}^q (\mathbf{c}^q)^T) \dots (\mathbf{I} - \mathbf{g}^{i+1} (\mathbf{c}^{i+1})^T) \mathbf{P}^i\end{aligned} \quad (7.8)$$

依次应用卡尔曼滤波校正方程式(7.3)、式(7.1)、式(7.8)和式(7.7), 有

$$\begin{aligned}\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \\ &= (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \hat{\mathbf{x}}_{k|k-1} \mathbf{G}_k \mathbf{v}_k \\ &= (\mathbf{I} - \mathbf{P}^q \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k) \hat{\mathbf{x}}_{k|k-1} + \mathbf{P}^q \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{v}_k \\ &= \left(\mathbf{I} - \sum_{i=1}^q \mathbf{P}^q \mathbf{c}^i (r^i)^{-1} (\mathbf{c}^i)^T \right) \hat{\mathbf{x}}^0 + \sum_{i=1}^q \mathbf{P}^q \mathbf{c}^i (r^i)^{-1} \mathbf{v}^i \\ &= \left[(\mathbf{I} - \mathbf{P}^q \mathbf{c}^q (r^q)^{-1} (\mathbf{c}^q)^T) - \sum_{i=1}^{q-1} (\mathbf{I} - \mathbf{g}^q (\mathbf{c}^q)^T) \right. \\ &\quad \left. \dots (\mathbf{I} - \mathbf{g}^{i+1} (\mathbf{c}^{i+1})^T) \mathbf{P}^i \mathbf{c}^i (r^i)^{-1} (\mathbf{c}^i)^T \right] \hat{\mathbf{x}}^0 + \sum_{i=1}^{q-1} (\mathbf{I} - \mathbf{g}^q (\mathbf{c}^q)^T) \\ &\quad \dots (\mathbf{I} - \mathbf{g}^{i+1} (\mathbf{c}^{i+1})^T) \mathbf{P}^i \mathbf{c}^i (r^i)^{-1} \mathbf{v}^i + \mathbf{P}^q \mathbf{c}^q (r^q)^{-1} \mathbf{v}^q \\ &= \left[\left(\mathbf{I} - \mathbf{g}^q (\mathbf{c}^q)^T \right) - \sum_{i=1}^{q-1} \left(\mathbf{I} - \mathbf{g}^q (\mathbf{c}^q)^T \right) \right. \\ &\quad \left. \dots (\mathbf{I} - \mathbf{g}^{i+1} (\mathbf{c}^{i+1})^T) \mathbf{g}^i (\mathbf{c}^i)^T \right] \hat{\mathbf{x}}^0 + \sum_{i=1}^{q-1} (\mathbf{I} - \mathbf{g}^q (\mathbf{c}^q)^T)\end{aligned}$$

$$\begin{aligned} & \cdots (\mathbf{I} - \mathbf{g}^{i+1}(\mathbf{c}^{i+1})^T) \mathbf{g}^i v^i + \mathbf{g}^q v^q \\ &= (\mathbf{I} - \mathbf{g}^q(\mathbf{c}^q)^T) \cdot \cdots \cdot (\mathbf{I} - \mathbf{g}^1(\mathbf{c}^1)^T) \hat{\mathbf{x}}_0 \\ &+ \sum_{i=1}^{q-1} (\mathbf{I} - \mathbf{g}^q(\mathbf{c}^q)^T) \cdot \cdots \cdot (\mathbf{I} - \mathbf{g}^{i+1}(\mathbf{c}^{i+1})^T) \mathbf{g}^i v^i + \mathbf{g}^q v^q \end{aligned}$$

另一方面,由式(7.2)中的第二个方程,可得

$$\begin{aligned} \hat{\mathbf{x}}^q &= (\mathbf{I} - \mathbf{g}^q(\mathbf{c}^q)^T) \hat{\mathbf{x}}^{q-1} + \mathbf{g}^q v^q \\ &= (\mathbf{I} - \mathbf{g}^q(\mathbf{c}^q)^T)(\mathbf{I} - \mathbf{g}^{q-1}(\mathbf{c}^{q-1})^T) \hat{\mathbf{x}}^{q-2} \\ &\quad + (\mathbf{I} - \mathbf{g}^q(\mathbf{c}^q)^T) \mathbf{g}^{q-1} v^{q-1} + \mathbf{g}^q v^q \\ &= \cdots \\ &= (\mathbf{I} - \mathbf{g}^q(\mathbf{c}^q)^T) \cdots (\mathbf{I} - \mathbf{g}^1(\mathbf{c}^1)^T) \hat{\mathbf{x}}^0 \\ &\quad + \sum_{i=1}^{q-1} (\mathbf{I} - \mathbf{g}^q(\mathbf{c}^q)^T) \cdots (\mathbf{I} - \mathbf{g}^{i+1}(\mathbf{c}^{i+1})^T) \mathbf{g}^i v^i + \mathbf{g}^q v^q \end{aligned}$$

该式与上面 $\hat{\mathbf{x}}_{k|k}$ 的表达式是相同的,即证明了 $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}^q$ 。

这就完成了定理 7.1 的证明。

7.2 平方根算法

下面介绍平方根算法。由线性代数理论得出的以下结论对该算法至关重要。下面引理由读者证明(见练习 7.1)。

引理 7.1 对于任意的正定对称矩阵 \mathbf{A} ,有唯一的下三角矩阵 \mathbf{A}^c ,使得 $\mathbf{A} = \mathbf{A}^c(\mathbf{A}^c)^T$ 。一般地,对于任意 $n \times (n+p)$ 矩阵 \mathbf{A} ,存在 $n \times n$ 矩阵 $\tilde{\mathbf{A}}$,使得 $\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T = \mathbf{A}\mathbf{A}^T$ 。

\mathbf{A}^c 的特性使其可以作为 \mathbf{A} 的平方根。又因为 \mathbf{A}^c 是下三角阵,因此它的逆可以很方便得到(见练习 7.3)。同样需要注意的是,在计算平方根时,非常小的数字会变得较大,非常大的数字会变得较小,因此计算会更精确。通过 Cholesky 因式分解(一种高斯消除方法),一个矩阵可以分解为一个下三角矩阵及其转置的乘积形式,这也是上标缩写为 c 的原因。对于一般情形, $\tilde{\mathbf{A}}$ 称作 $\mathbf{A}\mathbf{A}^T$ 的平方根。

在下面讨论的平方根算法中,对此下三角因式求逆:

$$\mathbf{H}_k := (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^c \quad (7.9)$$

为了提高算法的精度,我们用 \mathbf{R}_k^c 代替正定平方根 $\mathbf{R}_k^{1/2}$ 。当然,如果 \mathbf{R}_k 是对角阵,则有 $\mathbf{R}_k^c = \mathbf{R}_k^{1/2}$ 。

首先考虑下面的递推方案。令

$$\mathbf{J}_{0,0} = (\text{Var}(\mathbf{x}_0))^{1/2}$$

$\mathbf{J}_{k,k-1}$ 是下列矩阵的平方根:

$$[\mathbf{A}_{k-1} \mathbf{J}_{k-1,k-1} \quad \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}]_{n \times (n+p)} [\mathbf{A}_{k-1} \mathbf{J}_{k-1,k-1} \quad \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}]_{n \times (n+p)}^T$$

且有

$$\mathbf{J}_{k,k} = \mathbf{J}_{k,k-1} [\mathbf{I} - \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1}], \quad k = 1, 2, \dots$$

式中 $(\text{Var}(\mathbf{x}_0))^{1/2}$ 和 $\mathbf{Q}_{k-1}^{1/2}$ 分别是 $\text{Var}(\mathbf{x}_0)$ 和 \mathbf{Q}_{k-1} 的任意平方根。辅助矩阵 $\mathbf{J}_{k,k-1}$ 和 $\mathbf{J}_{k,k}$ 尽管不一定是下三角阵或者正定阵,但它们也是(分别是 $\mathbf{P}_{k,k-1}$ 和 $\mathbf{P}_{k,k}$ 的)平方根,即有下述定理。

定理 7.2 $\mathbf{J}_{0,0} \mathbf{J}_{0,0}^T = \mathbf{P}_{0,0}$,且对于 $k=1,2,\dots$,有

$$\mathbf{J}_{k,k-1} \mathbf{J}_{k,k-1}^T = \mathbf{P}_{k,k-1} \quad (7.10)$$

$$\mathbf{J}_{k,k} \mathbf{J}_{k,k}^T = \mathbf{P}_{k,k} \quad (7.11) \blacksquare$$

因为 $\mathbf{P}_{0,0} = \text{Var}(\mathbf{x}_0)$,所以定理中第一个结论显然成立。通过数学归纳法可以证明式(7.10)和式(7.11)。假定对于 $k-1$,式(7.11)成立;然后利用卡尔曼滤波过程中 $\mathbf{P}_{k,k-1}$ 和 $\mathbf{P}_{k-1,k-1}$ 之间的关系,就能立即得到式(7.10)。对于同样的 k ,可以用式(7.10)验证式(7.11)。

由于

$$\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T = \mathbf{H}_k \mathbf{H}_k^T - \mathbf{R}_k$$

有

$$\begin{aligned} & (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} + [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \\ & - (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \\ & = (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \{ \mathbf{H}_k (\mathbf{H}_k + \mathbf{R}_k^c)^T + (\mathbf{H}_k + \mathbf{R}_k^c) \mathbf{H}_k^T - \mathbf{H}_k \mathbf{H}_k^T + \mathbf{R}_k \} \\ & \cdot [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \\ & = (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \{ \mathbf{H}_k \mathbf{H}_k^T + \mathbf{H}_k (\mathbf{R}_k^c)^T + \mathbf{R}_k^c \mathbf{H}_k^T + \mathbf{R}_k \} \\ & \cdot [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \\ & = (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c) (\mathbf{H}_k + \mathbf{R}_k^c)^T [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \\ & = (\mathbf{H}_k^T)^{-1} \mathbf{H}_k^{-1} \\ & = (\mathbf{H}_k \mathbf{H}_k^T)^{-1} \end{aligned}$$

由式(7.10)可得

$$\begin{aligned} \mathbf{J}_{k,k} \mathbf{J}_{k,k}^T &= \mathbf{J}_{k,k-1} [\mathbf{I} - \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1}] \\ &\quad \cdot [\mathbf{I} - \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1}] \mathbf{J}_{k,k-1}^T \\ &= \mathbf{J}_{k,k-1} \{ \mathbf{I} - \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1} \\ &\quad - \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1} \\ &\quad + \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1} \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T \} \end{aligned}$$

$$\begin{aligned}
& \cdot [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1} \} \mathbf{J}_{k,k-1}^T \\
& = \mathbf{P}_{k,k-1} - \mathbf{P}_{k,k-1} \mathbf{C}_k^T [(\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} + [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1} \\
& \quad - (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T [(\mathbf{H}_k + \mathbf{R}_k^c)^T]^{-1} \mathbf{H}_k^{-1}] \mathbf{C}_k \mathbf{P}_{k,k-1} \\
& = \mathbf{P}_{k,k-1} - \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{H}_k \mathbf{H}_k^T)^{-1} \mathbf{C}_k \mathbf{P}_{k,k-1} \\
& = \mathbf{P}_{k,k}
\end{aligned}$$

至此,归纳过程完成。

总的来说,平方根卡尔曼滤波算法归纳如下:

(i) 计算 $\mathbf{J}_{0,0} = (\text{Var}(\mathbf{x}_0))^{1/2}$ 。

(ii) 对于 $k=1, 2, \dots$, 计算 $\mathbf{J}_{k,k-1}$, 它是

$$[\mathbf{A}_{k-1} \mathbf{J}_{k-1,k-1} \quad \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}]_{n \times (n+p)} [\mathbf{A}_{k-1} \mathbf{J}_{k-1,k-1} \quad \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}]_{n \times (n+p)}^T$$

的平方根,且设置矩阵

$$\mathbf{H}_k = (\mathbf{C}_k \mathbf{J}_{k,k-1} \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T + \mathbf{R}_k)^c$$

然后计算

$$\mathbf{J}_{k,k} = \mathbf{J}_{k,k-1} [\mathbf{I} - \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1}]$$

(iii) 对于 $k=1, 2, \dots$, 令 $\hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0)$, 并由(ii)得到的结论, 计算

$$\mathbf{G}_k = \mathbf{J}_{k,k-1} \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} \mathbf{H}_k^{-1}$$

和

$$\hat{\mathbf{x}}_{k|k} = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1})$$

(见图 7.2)。

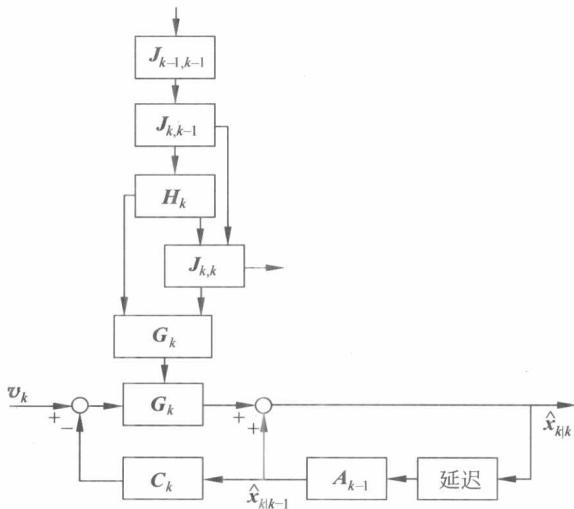


图 7.2 平方根卡尔曼滤波算法

这里注意到,只需对三角阵进行求逆运算。此外,这些三角矩阵是另外一些矩阵的平方根,而这些矩阵可能有非常大或非常小的元素。

7.3 实时应用算法

特殊情况下,当

$$\mathbf{R}_k = \text{diag}[r_k^1, \dots, r_k^q]$$

是对角阵时,序贯算法和平方根算法可以融合成下面的新算法。这种新算法不需要对矩阵直接求逆。

- (i) 计算 $\mathbf{J}_{0,0} = (\text{Var}(\mathbf{x}_0))^{1/2}$ 。
- (ii) 对于每一个固定的 $k=1, 2, \dots$, 计算
- (a) $\mathbf{J}_{k,k-1}$, 它是矩阵

$$[\mathbf{A}_{k-1} \mathbf{J}_{k-1,k-1} \quad \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}]_{n \times (n+p)} \quad [\mathbf{A}_{k-1} \mathbf{J}_{k-1,k-1} \quad \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}]_{n \times (n+p)}^T$$

的平方根。

- (b) 对于 $i=1, 2, \dots, k$,

$$\begin{aligned} \mathbf{g}_k^i &= \frac{1}{(\mathbf{c}_k^i)^T \mathbf{J}_{k,k-1}^{i-1} (\mathbf{J}_{k,k-1}^{i-1})^T \mathbf{c}_k^i + r_k^i} \mathbf{J}_{k,k-1}^{i-1} (\mathbf{J}_{k,k-1}^{i-1})^T \mathbf{c}_k^i, \\ \mathbf{J}_{k,k-1}^i &= (\mathbf{J}_{k,k-1}^{i-1} (\mathbf{J}_{k,k-1}^{i-1})^T - \mathbf{g}_k^i (\mathbf{c}_k^i)^T \mathbf{J}_{k,k-1}^{i-1} (\mathbf{J}_{k,k-1}^{i-1})^T)^c \end{aligned}$$

式中 $\mathbf{J}_{k,k-1}^0 := \mathbf{J}_{k,k-1}$, $\mathbf{J}_{k,k-1}^q = \mathbf{J}_{k,k}$, $\mathbf{C}_k^T := [\mathbf{c}_k^1 \cdots \mathbf{c}_k^q]$ 。

- (iii) 计算 $\hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0)$ 。
- (iv) 对于每一个固定的 $k=1, 2, \dots$, 计算
- (a) $\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}$ 。

(b) 对于 $i=1, 2, \dots, q$, 且 $\hat{\mathbf{x}}_k^0 := \hat{\mathbf{x}}_{k|k-1}$, 并利用 (ii)(b) 中的信息, 计算

$$\hat{\mathbf{x}}_k^i = \hat{\mathbf{x}}_k^{i-1} + (v_k^i - (\mathbf{c}_k^i)^T \hat{\mathbf{x}}_k^{i-1}) \mathbf{g}_k^i$$

式中 $v_k := [v_k^1 \cdots v_k^q]^T$, 所以有

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_k^q$$

练习

7.1 给出引理 7.1 的证明。

7.2 寻找下三角矩阵 L , 使其满足:

$$(a) \mathbf{L}\mathbf{L}^T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 8 & 2 \\ 3 & 2 & 14 \end{bmatrix}$$

$$(b) \mathbf{L}\mathbf{L}^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 4 \end{bmatrix}$$

7.3 (a) 推导下面的矩阵求逆公式: $\mathbf{L} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$, 式中 l_{11}, l_{22} 和 l_{33}

非零。

(b) 给出下面矩阵逆的表达式:

$$\mathbf{L}^{-1} = \begin{bmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ l_{n1} & l_{n2} & \cdots & \cdots & l_{nn} \end{bmatrix}$$

式中 l_{11}, \dots, l_{nn} 非零。

7.4 考虑卡尔曼滤波过程的计算机仿真。令 $\epsilon \ll 1$ 是很小的正数, 且

$$1 - \epsilon \not\approx 1$$

$$1 - \epsilon^2 \approx 1$$

式中“ \approx ”表示经过四舍五入之后的相等。假设有 $\mathbf{P}_{k,k} = \begin{bmatrix} \frac{\epsilon^2}{1-\epsilon^2} & 0 \\ 0 & 1 \end{bmatrix}$ 。对此例,

比较标准卡尔曼滤波和平方根滤波。注意, 此例说明平方根滤波改进了数值特性。

7.5 证明: 对于任意的正定对称阵 \mathbf{A} , 存在唯一的上三角阵 \mathbf{A}^u , 使得 $\mathbf{A} = \mathbf{A}^u (\mathbf{A}^u)^T$ 。

7.6 用上三角分解代替下三角分解, 推导出新的平方根卡尔曼滤波。

7.7 结合序贯算法和上三角分解的平方根算法, 推导出新的卡尔曼滤波算法。

第 8 章

扩展卡尔曼滤波和系统辨识

设计卡尔曼滤波过程是为了估计线性模型中的状态向量。如果模型是非线性的，通常在推导滤波方程时，增加线性化步骤。在状态估计时，对系统方程在前一状态估计值处做实时的线性泰勒近似；在预测步骤中，对量测方程在相应的预测位置也进行线性泰勒近似；所得到的卡尔曼滤波被称为扩展卡尔曼滤波。处理非线性模型的这一思想是很自然的，且滤波过程简单有效。此外，它有许多重要的实时应用，其中的一个应用是自适应系统辨识，本章对此也进行了简单介绍。最后，对卡尔曼滤波算法的线性化步骤进行改进，介绍了一种改进扩展卡尔曼滤波方案，该方案拥有并行的计算结构；并给出了两个数值例子来证明改进卡尔曼滤波在状态估计和系统参数辨识两方面都优于标准卡尔曼滤波。

8.1 扩展卡尔曼滤波

一个不一定是线性系统的状态空间描述称为是非线性系统模型。本章将研究形如下面的非线性模型：

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{H}_k(\mathbf{x}_k) \xi_k \\ \mathbf{v}_k = \mathbf{g}_k(\mathbf{x}_k) + \eta_k \end{cases} \quad (8.1)$$

式中 \mathbf{f}_k 和 \mathbf{g}_k 分别是 \mathbf{R}^n 和 \mathbf{R}^q 上的向量值函数， $1 \leq q \leq n$ ， \mathbf{H}_k 是 $\mathbf{R}^n \times \mathbf{R}^q$ 上的矩阵值函数，对于每个 k ， $\mathbf{f}_k(\mathbf{x}_k)$ 和 $\mathbf{g}_k(\mathbf{x}_k)$ 对 \mathbf{x}_k 所有分量的一阶偏导都是连续的。如常，设 $\{\xi_k\}$ 和 $\{\eta_k\}$ 为 \mathbf{R}^p 和 \mathbf{R}^q 上的零均值高斯白噪声序列，其中 $1 \leq p, q \leq n$ ，且对于所有的 k 和 l ，有

$$E(\underline{\xi}_k \underline{\xi}_l^T) = Q_k \delta_{kl}, \quad E(\underline{\eta}_k \underline{\eta}_l^T) = R_k \delta_{kl}$$

$$E(\underline{\xi}_k \underline{\eta}_l^T) = \mathbf{0}, \quad E(\underline{\xi}_k \underline{x}_0^T) = \mathbf{0}, \quad E(\underline{\eta}_k \underline{x}_0^T) = \mathbf{0}$$

实时线性化过程如下。

为了与线性模型保持一致,初始估计 $\hat{x}_0 = \hat{x}_{0|0}$ 和预测值 $\hat{x}_{1|0}$ 选择为

$$\hat{x}_0 = E(x_0), \quad \hat{x}_{1|0} = f_0(\hat{x}_0)$$

应用式 $\hat{x}_k = \hat{x}_{k|k}$, 对于 $k=1, 2, \dots$, 依次使用预测值, 可得

$$\hat{x}_{k+1|k} = f_k(\hat{x}_k) \quad (8.2)$$

及线性状态空间描述:

$$\begin{cases} \mathbf{x}_{k+1} = A_k \mathbf{x}_k + u_k + \Gamma_k \underline{\xi}_k \\ w_k = C_k \mathbf{x}_k + \underline{\eta}_k \end{cases} \quad (8.3)$$

式中, A_k 、 u_k 、 Γ_k 、 w_k 和 C_k 能够如以下计算实时求出。

对于 $j=0, 1, \dots, k$, 假定 \hat{x}_j 能够求出, 那么 $\hat{x}_{j+1|j}$ 也能够用式(8.2)定义。考虑 $f_k(x_k)$ 在 \hat{x}_k 处的线性泰勒近似和 $g_k(x_k)$ 在 $\hat{x}_{k|k-1}$ 处的线性泰勒近似:

$$\begin{cases} f_k(x_k) \approx f_k(\hat{x}_k) + A_k(x_k - \hat{x}_k) \\ g_k(x_k) \approx g_k(\hat{x}_{k|k-1}) + C_k(x_k - \hat{x}_{k|k-1}) \end{cases} \quad (8.4)$$

式中

$$A_k = \left[\frac{\partial f_k}{\partial x_k}(\hat{x}_k) \right], \quad C_k = \left[\frac{\partial g_k}{\partial x_k}(\hat{x}_{k|k-1}) \right] \quad (8.5)$$

此处及往后, 对于任何向量值函数

$$\mathbf{h}(x_k) = \begin{bmatrix} h_1(x_k) \\ \vdots \\ h_m(x_k) \end{bmatrix}$$

式中

$$x_k = \begin{bmatrix} x_k^1 \\ \vdots \\ x_k^n \end{bmatrix}$$

令

$$\left[\frac{\partial \mathbf{h}}{\partial x_k}(x_k^*) \right] = \begin{bmatrix} \frac{\partial h_1}{\partial x_k^1}(x_k^*) & \cdots & \frac{\partial h_1}{\partial x_k^n}(x_k^*) \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_k^1}(x_k^*) & \cdots & \frac{\partial h_m}{\partial x_k^n}(x_k^*) \end{bmatrix} \quad (8.6)$$

并设置

$$\begin{cases} \mathbf{u}_k = f_k(\hat{\mathbf{x}}_k) - A_k \hat{\mathbf{x}}_k \\ \mathbf{F}_k = H_k(\hat{\mathbf{x}}_k) \\ \mathbf{w}_k = \mathbf{v}_k - g_k(\hat{\mathbf{x}}_{k|k-1}) + C_k \hat{\mathbf{x}}_{k|k-1} \end{cases} \quad (8.7)$$

在 k 时刻, 使用式(8.5)和式(8.7)中定义的矩阵和向量, 非线性模型(8.1)可以由线性模型(8.3)近似(见练习 8.2)。当然, 只有 $\hat{\mathbf{x}}_k$ 能够确定, 线性化才能实现。我们已经获得 $\hat{\mathbf{x}}_0$, 所以可以得到 $k=0$ 时(8.3)的系统方程。那么使用数据 $[\mathbf{v}_0^T \quad \mathbf{w}_1^T]^T$, 定义 $\hat{\mathbf{x}}_1 = \hat{\mathbf{x}}_{1|1}$ 作为线性模型(8.3)中 \mathbf{x}_1 的最优无偏估计(采用最优权重)。则对 $k=1$, 根据式(8.2), 得到式(8.3)。同理类似地使用数据 $[\mathbf{v}_0^T \quad \mathbf{w}_1^T \quad \mathbf{w}_2^T]^T$, 可以得到 $\hat{\mathbf{x}}_2 = \hat{\mathbf{x}}_{2|2}$ 。从第 2 章和第 3 章中线性确定性状态空间描述的卡尔曼滤波结果(见式(2.18)和练习 3.6), 可以得到“校正的”公式:

$$\begin{aligned} \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k|k-1} + G_k(w_k - C_k \hat{\mathbf{x}}_{k|k-1}) \\ &= \hat{\mathbf{x}}_{k|k-1} + G_k((\mathbf{v}_k - g_k(\hat{\mathbf{x}}_{k|k-1}) + C_k \hat{\mathbf{x}}_{k|k-1}) - C_k \hat{\mathbf{x}}_{k|k-1}) \\ &= \hat{\mathbf{x}}_{k|k-1} + G_k(\mathbf{v}_k - g_k(\hat{\mathbf{x}}_{k|k-1})) \end{aligned}$$

式中 G_k 是 k 时刻线性模型(8.3)中的卡尔曼增益矩阵。

最终的滤波过程称为扩展卡尔曼滤波。滤波算法总结如下(见练习 8.3):

$$\begin{cases} P_{0,0} = \text{Var}(\mathbf{x}_0) \\ \hat{\mathbf{x}}_0 = E(\mathbf{x}_0) \\ \text{对于 } k = 1, 2, \dots \\ P_{k,k-1} = \left[\frac{\partial f_{k-1}}{\partial \mathbf{x}_{k-1}}(\hat{\mathbf{x}}_{k-1}) \right] P_{k-1,k-1} \left[\frac{\partial f_{k-1}}{\partial \mathbf{x}_{k-1}}(\hat{\mathbf{x}}_{k-1}) \right]^T + H_{k-1}(\hat{\mathbf{x}}_{k-1}) Q_{k-1} H_{k-1}^T(\hat{\mathbf{x}}_{k-1}) \\ \hat{\mathbf{x}}_{k|k-1} = f_{k-1}(\hat{\mathbf{x}}_{k-1}) \\ G_k = P_{k,k-1} \left[\frac{\partial g_k}{\partial \mathbf{x}_k}(\hat{\mathbf{x}}_{k|k-1}) \right]^T \cdot \left[\left[\frac{\partial g_k}{\partial \mathbf{x}_k}(\hat{\mathbf{x}}_{k|k-1}) \right] P_{k,k-1} \left[\frac{\partial g_k}{\partial \mathbf{x}_k}(\hat{\mathbf{x}}_{k|k-1}) \right]^T + R_k \right]^{-1} \\ P_{k,k} = \left[I - G_k \left[\frac{\partial g_k}{\partial \mathbf{x}_k}(\hat{\mathbf{x}}_{k|k-1}) \right] \right] P_{k,k-1} \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + G_k(\mathbf{v}_k - g_k(\hat{\mathbf{x}}_{k|k-1})) \end{cases} \quad (8.8)$$

(见图 8.1)。

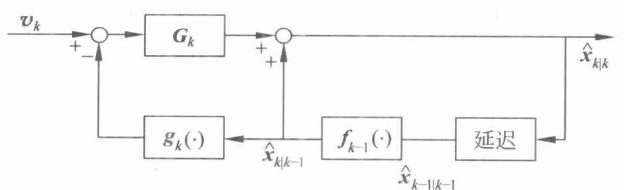


图 8.1 扩展卡尔曼滤波算法

8.2 卫星轨道估计

卫星的平面轨道估计给扩展卡尔曼滤波提供了一个有趣的例子。卫星在平面轨道上的运动控制方程为

$$\begin{cases} \ddot{r} = r\dot{\theta}^2 - mg r^{-2} + \xi_r \\ \ddot{\theta} = -2r^{-1}\dot{r}\dot{\theta} + r^{-1}\xi_\theta \end{cases} \quad (8.9)$$

式中 r 为卫星距离地球中心(被称作吸引点)的(径向)距离; θ 是以某个坐标轴为参考所测量的角度(被称作角位移); m, g 为常数, 分别是地球质量和万有引力常数(见图 8.2)。此外, ξ_r 和 ξ_θ 假设为连续时间不相关的零均值高斯白噪声过程。令

$$\mathbf{x} = [r \quad \dot{r} \quad \theta \quad \dot{\theta}]^T = [\mathbf{x}[1] \quad \mathbf{x}[2] \quad \mathbf{x}[3] \quad \mathbf{x}[4]]^T$$

则式(8.9)的方程变为

$$\dot{\mathbf{x}} = \begin{bmatrix} \mathbf{x}[2] \\ \mathbf{x}[1]\mathbf{x}[4]^2 - mg/\mathbf{x}[1]^2 \\ \mathbf{x}[4] \\ -2\mathbf{x}[2]\mathbf{x}[4]/\mathbf{x}[1] \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/\mathbf{x}[1] \end{bmatrix} \begin{bmatrix} \xi_r \\ 0 \\ 0 \\ \xi_\theta \end{bmatrix}$$

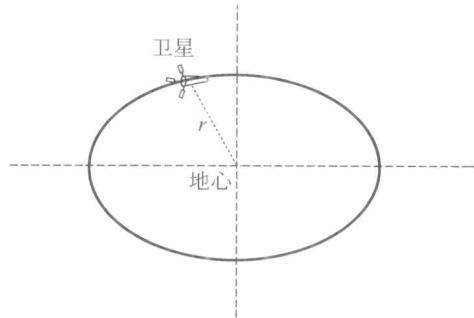


图 8.2 卫星轨道示意图

用 \mathbf{x}_k 代替 \mathbf{x} , 用 $(\mathbf{x}_{k+1} - \mathbf{x}_k)h^{-1}$ 代替 $\dot{\mathbf{x}}$, 其中 $h > 0$ 表示采样时间, 得到离散非线性模型:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k) + \mathbf{H}(\mathbf{x}_k) \underline{\xi}_k$$

式中

$$\begin{aligned} \mathbf{f}(\mathbf{x}_k) &= \begin{bmatrix} \mathbf{x}_k[1] + h\mathbf{x}_k[2] \\ \mathbf{x}_k[2] + h\mathbf{x}_k[1]\mathbf{x}_k[4]^2 - hmg/\mathbf{x}_k[1]^2 \\ \mathbf{x}_k[3] + h\mathbf{x}_k[4] \\ \mathbf{x}_k[4] - 2h\mathbf{x}_k[2]\mathbf{x}_k[4]/\mathbf{x}_k[1] \end{bmatrix} \\ \mathbf{H}(\mathbf{x}_k) &= \begin{bmatrix} h & 0 & 0 & 0 \\ 0 & h & 0 & 0 \\ 0 & 0 & h & 0 \\ 0 & 0 & 0 & h/\mathbf{x}_k[1] \end{bmatrix} \\ \boldsymbol{\xi}_k &:= [0 \quad \xi_r(k) \quad 0 \quad \xi_\theta(k)]^T \end{aligned}$$

假设距离 r 已由测量得知, 给出数据信息:

$$\mathbf{v}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}_k + \boldsymbol{\eta}_k$$

式中 $\{\boldsymbol{\eta}_k\}$ 是零均值高斯白噪声序列, 与 $\{\xi_r(k)\}$ 和 $\{\xi_\theta(k)\}$ 是相互独立的。则

$$\begin{aligned} &\left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}_{k-1}}(\hat{\mathbf{x}}_{k-1}) \right] \\ &= \begin{bmatrix} 1 & 2hmg/\hat{\mathbf{x}}_{k-1}[1]^3 + h\hat{\mathbf{x}}_{k-1}[4]^2 & 0 & 2h\hat{\mathbf{x}}_{k-1}[2]\hat{\mathbf{x}}_{k-1}[4]/\hat{\mathbf{x}}_{k-1}[1]^2 \\ h & 1 & 0 & -2h\hat{\mathbf{x}}_{k-1}[4]/\hat{\mathbf{x}}_{k-1}[1] \\ 0 & 0 & 1 & 0 \\ 0 & 2h\hat{\mathbf{x}}_{k-1}[1]\hat{\mathbf{x}}_{k-1}[4] & h & 1 - 2h\hat{\mathbf{x}}_{k-1}[2]/\hat{\mathbf{x}}_{k-1}[1] \end{bmatrix}^T \\ \mathbf{H}(\hat{\mathbf{x}}_{k-1|k-1}) &= \begin{bmatrix} h & 0 & 0 & 0 \\ 0 & h & 0 & 0 \\ 0 & 0 & h & 0 \\ 0 & 0 & 0 & h/\hat{\mathbf{x}}_{k-1}[1] \end{bmatrix} \\ \left[\frac{\partial \mathbf{g}}{\partial \mathbf{x}_k}(\hat{\mathbf{x}}_{k|k-1}) \right] &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ \mathbf{g}(\hat{\mathbf{x}}_{k|k-1}) &= \hat{\mathbf{x}}_{k|k-1}[1] \end{aligned}$$

通过使用扩展卡尔曼滤波方程(8.8)中的这些变量, 可以得到确定 $\hat{\mathbf{x}}_k$ 的算法, 据此可以估计出卫星平面轨道($\hat{\mathbf{x}}_k[1]$ 、 $\hat{\mathbf{x}}_k[3]$)。

8.3 自适应系统辨识

作为扩展卡尔曼滤波的一个应用, 下面我们讨论自适应系统辨识的问题。假设一个线性系统的状态空间描述为

$$\begin{cases} \underline{x}_{k+1} = A_k(\underline{\theta}) \underline{x}_k + \Gamma_k(\underline{\theta}) \underline{\xi}_k \\ \underline{v}_k = C_k(\underline{\theta}) \underline{x}_k + \underline{\eta}_k \end{cases} \quad (8.10)$$

式中, $\underline{x}_k \in \mathbf{R}^n$, $\underline{\xi}_k \in \mathbf{R}^p$, $\underline{\eta}_k \in \mathbf{R}^q$, $1 \leq p, q \leq n$, $\{\underline{\xi}_k\}$ 和 $\{\underline{\eta}_k\}$ 是不相关的高斯白噪声序列。在这个应用中, 假定 $A_k(\underline{\theta})$ 、 $\Gamma_k(\underline{\theta})$ 和 $C_k(\underline{\theta})$ 是某个未知常向量 $\underline{\theta}$ 的已知矩阵值函数。目标是“辨识” $\underline{\theta}$ 。

因为 $\underline{\theta}$ 是一个常值向量, 故此应该有 $\underline{\theta}_{k+1} = \underline{\theta}_k = \underline{\theta}$ 。但是这个假设并不能帮助我们解决任何问题, 这一点我们在下面及下节的简单例子中也会发现。事实上, $\underline{\theta}$ 必须被当作一个随机常向量, 例如:

$$\underline{\theta}_{k+1} = \underline{\theta}_k + \underline{\zeta}_k \quad (8.11)$$

式中 $\{\underline{\zeta}_k\}$ 是任意零均值高斯白噪声序列, 并与 $\{\underline{\eta}_k\}$ 不相关, 且其方差可以被预先设置成正定阵 $\text{Var}(\underline{\zeta}_k) = S_k$ 。在应用中, 对于所有 k , 可以选择 $S_k = S > \mathbf{0}$ (参考 8.4 节)。现在, 具备假设(8.11)的系统(8.10)可以改写为非线性模型:

$$\begin{cases} \begin{bmatrix} \underline{x}_{k+1} \\ \underline{\theta}_{k+1} \end{bmatrix} = \begin{bmatrix} A_k(\underline{\theta}_k) \underline{x}_k \\ \underline{\theta}_k \end{bmatrix} + \begin{bmatrix} \Gamma_k(\underline{\theta}_k) \underline{\xi}_k \\ \underline{\zeta}_k \end{bmatrix} \\ \underline{v}_k = [C_k(\underline{\theta}_k) \quad \mathbf{0}] \begin{bmatrix} \underline{x}_k \\ \underline{\theta}_k \end{bmatrix} + \underline{\eta}_k \end{cases} \quad (8.12)$$

扩展卡尔曼滤波可以用来估计状态向量, $\underline{\theta}_k$ 是该向量的分量。也就是说, $\underline{\theta}_k$ 以一种自适应的方式被最优估计。为了应用卡尔曼滤波过程(8.8), 仍需初始化 $\hat{\underline{\theta}}_0 := \hat{\underline{\theta}}_{0|0}$ 。一种方法是使用状态空间描述(8.10)。例如, 由于 $E(\underline{v}_0) = C_0(\underline{\theta})E(\underline{x}_0)$, 则 $\underline{v}_0 - C_0(\underline{\theta})E(\underline{x}_0)$ 均值为零, 从而可以从 $k=0$ 开始, 求出下面改进“观测方程”两侧的方差:

$$\underline{v}_0 - C_0(\underline{\theta})E(\underline{x}_0) = C_0(\underline{\theta})\underline{x}_0 - C_0(\underline{\theta})E(\underline{x}_0) + \underline{\eta}_0$$

并使用 $\text{Var}(\underline{v}_0 - C_0(\underline{\theta})E(\underline{x}_0))$ 的估计值 $[\underline{v}_0 - C_0(\underline{\theta})E(\underline{x}_0)][\underline{v}_0 - C_0(\underline{\theta})E(\underline{x}_0)]^\top$ (见练习 2.12), 近似得到

$$\begin{aligned} \underline{v}_0 \underline{v}_0^\top - C_0(\underline{\theta})E(\underline{x}_0) \underline{v}_0^\top - \underline{v}_0 (C_0(\underline{\theta})E(\underline{x}_0))^\top \\ + C_0(\underline{\theta})(E(\underline{x}_0)E(\underline{x}_0^\top) - \text{Var}(\underline{x}_0))C_0^\top(\underline{\theta}) - R_0 = \mathbf{0} \end{aligned} \quad (8.13)$$

(见练习 8.4)。现在来求解 $\underline{\theta}$, 并设置“最合适”的一个解作为初始估计 $\hat{\underline{\theta}}_0$ 。如果(8.13)关于 $\underline{\theta}$ 无解, 则可以借助方程

$$\begin{aligned}\boldsymbol{v}_1 &= \mathbf{C}_1(\underline{\boldsymbol{\theta}}) \boldsymbol{x}_1 + \underline{\boldsymbol{\eta}}_1 \\ &= \mathbf{C}_1(\underline{\boldsymbol{\theta}})(\mathbf{A}_0(\underline{\boldsymbol{\theta}}) \boldsymbol{x}_0 + \boldsymbol{\Gamma}_0(\underline{\boldsymbol{\theta}}) \underline{\boldsymbol{\xi}}_0) + \underline{\boldsymbol{\eta}}_1\end{aligned}$$

并使用相同的步骤,近似得到

$$\begin{aligned}&\boldsymbol{v}_1 \boldsymbol{v}_1^T - \mathbf{C}_1(\underline{\boldsymbol{\theta}}) \mathbf{A}_0(\underline{\boldsymbol{\theta}}) \mathbf{E}(\boldsymbol{x}_0) \boldsymbol{v}_1^T - \boldsymbol{v}_1 (\mathbf{C}_1(\underline{\boldsymbol{\theta}}) \mathbf{A}_0(\underline{\boldsymbol{\theta}}) \mathbf{E}(\boldsymbol{x}_0))^T \\ &- \mathbf{C}_1(\underline{\boldsymbol{\theta}}) \boldsymbol{\Gamma}_0(\underline{\boldsymbol{\theta}}) \mathbf{Q}_0 (\mathbf{C}_1(\underline{\boldsymbol{\theta}}) \boldsymbol{\Gamma}_0(\underline{\boldsymbol{\theta}}))^T + \mathbf{C}_1(\underline{\boldsymbol{\theta}}) \mathbf{A}_0(\underline{\boldsymbol{\theta}}) [\mathbf{E}(\boldsymbol{x}_0) \mathbf{E}(\boldsymbol{x}_0^T)] \\ &- \text{Var}(\boldsymbol{x}_0) [\mathbf{A}_0^T(\underline{\boldsymbol{\theta}}) \mathbf{C}_1^T(\underline{\boldsymbol{\theta}}) - \mathbf{R}_1] = \mathbf{0}\end{aligned}\quad (8.14)$$

(见练习 8.5)。

一旦选定 $\hat{\boldsymbol{\theta}}_0$, 应用扩展卡尔曼滤波过程(8.8),便得到以下算法:

$$\left\{ \begin{array}{l} \begin{aligned} \begin{bmatrix} \hat{\boldsymbol{x}}_0 \\ \hat{\boldsymbol{\theta}}_0 \end{bmatrix} &= \begin{bmatrix} \mathbf{E}(\boldsymbol{x}_0) \\ \hat{\boldsymbol{\theta}}_0 \end{bmatrix} \\ \mathbf{P}_{0,0} &= \begin{bmatrix} \text{Var}(\boldsymbol{x}_0) & \mathbf{0} \\ \mathbf{0} & S_0 \end{bmatrix} \end{aligned} \\ \text{对于 } k = 1, 2, \dots, \text{计算} \\ \begin{aligned} \begin{bmatrix} \hat{\boldsymbol{x}}_{k|k-1} \\ \hat{\boldsymbol{\theta}}_{k|k-1} \end{bmatrix} &= \begin{bmatrix} \mathbf{A}_{k-1}(\hat{\boldsymbol{\theta}}_{k-1}) \hat{\boldsymbol{x}}_{k-1} \\ \hat{\boldsymbol{\theta}}_{k-1} \end{bmatrix} \\ \mathbf{P}_{k,k-1} &= \begin{bmatrix} \mathbf{A}_{k-1}(\hat{\boldsymbol{\theta}}_{k-1}) & \frac{\partial}{\partial \underline{\boldsymbol{\theta}}} [\mathbf{A}_{k-1}(\hat{\boldsymbol{\theta}}_{k-1}) \hat{\boldsymbol{x}}_{k-1}] \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{P}_{k-1,k-1} \\ &\cdot \begin{bmatrix} \mathbf{A}_{k-1}(\hat{\boldsymbol{\theta}}_{k-1}) & \frac{\partial}{\partial \underline{\boldsymbol{\theta}}} [\mathbf{A}_{k-1}(\hat{\boldsymbol{\theta}}_{k-1}) \hat{\boldsymbol{x}}_{k-1}] \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^T \\ &+ \begin{bmatrix} \boldsymbol{\Gamma}_{k-1}(\hat{\boldsymbol{\theta}}_{k-1}) \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^T(\hat{\boldsymbol{\theta}}_{k-1}) & \mathbf{0} \\ \mathbf{0} & S_{k-1} \end{bmatrix} \\ \mathbf{G}_k &= \mathbf{P}_{k,k-1} \left[\mathbf{C}_k(\hat{\boldsymbol{\theta}}_{k|k-1}) \quad \mathbf{0} \right]^T \cdot \left[\left[\mathbf{C}_k(\hat{\boldsymbol{\theta}}_{k|k-1}) \quad \mathbf{0} \right] \mathbf{P}_{k,k-1} \left[\mathbf{C}_k(\hat{\boldsymbol{\theta}}_{k|k-1}) \quad \mathbf{0} \right]^T + \mathbf{R}_k \right]^{-1} \\ \mathbf{P}_{k,k} &= \left[\mathbf{I} - \mathbf{G}_k \left[\mathbf{C}_k(\hat{\boldsymbol{\theta}}_{k|k-1}) \quad \mathbf{0} \right] \right] \mathbf{P}_{k,k-1} \\ \begin{bmatrix} \hat{\boldsymbol{x}}_k \\ \hat{\boldsymbol{\theta}}_k \end{bmatrix} &= \begin{bmatrix} \hat{\boldsymbol{x}}_{k|k-1} \\ \hat{\boldsymbol{\theta}}_{k|k-1} \end{bmatrix} + \mathbf{G}_k (\boldsymbol{v}_k - \mathbf{C}_k(\hat{\boldsymbol{\theta}}_{k|k-1}) \hat{\boldsymbol{x}}_{k|k-1}) \end{array} \right]\quad (8.15)$$

(见练习 8.6)。

我们注意到,如果未知常向量 $\underline{\theta}$ 是确定的,也就是说 $\underline{\theta}_{k+1} = \underline{\theta}_k = \underline{\theta}$,因而有 $\mathbf{S}_k = \mathbf{0}$,然后对于所有 k ,步骤(8.15)只能得到与观测数据无关的 $\hat{\underline{\theta}}_k = \hat{\underline{\theta}}_{k-1}$ (见练习8.7),并不能给出关于 $\hat{\underline{\theta}}_k$ 的任何信息。换句话说,如果 $\mathbf{S}_k = \mathbf{0}$,则未知系统参数向量 $\underline{\theta}$ 不能由扩展卡尔曼滤波技术辨识出来。

8.4 一个常值参数辨识的例子^①

下面的简单例子能够说明扩展卡尔曼滤波如何能够很好地完成自适应系统辨识,甚至是在初始估计 $\hat{\underline{\theta}}_0$ 是任意选取的情况下。

考虑一个线性系统,其状态空间描述为

$$\begin{cases} x_{k+1} = ax_k \\ v_k = x_k + \eta_k \end{cases}$$

式中 a 是需要辨识的未知参数。现在,将 a 作为一个随机变量;即,我们认为:

$$a_{k+1} = a_k + \zeta_k$$

式中 a_k 是 a 在 k 时刻的值, $E(\zeta_k) = 0$, $\text{Var}(\zeta_k) = 0.01$ 。假如 $E(x_0) = 1$, $\text{Var}(x_0) = 0.01$, $\{\eta_k\}$ 是零均值高斯白噪声序列,且有 $\text{Var}(\eta_k) = 0.01$ 。我们的目标是通过执行卡尔曼滤波,来估计未知参数 a_k 。在系统方程中,用 a_k 替代 a ,那么上面的方程可变为下面的非线性状态空间描述:

$$\begin{cases} \begin{bmatrix} x_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} a_k x_k \\ a_k \end{bmatrix} + \begin{bmatrix} 0 \\ \zeta_k \end{bmatrix} \\ v_k = [1 \quad 0] \begin{bmatrix} x_k \\ a_k \end{bmatrix} + \eta_k \end{cases} \quad (8.16)$$

对该模型应用(8.15)可得

$$\left\{ \begin{array}{l} \mathbf{P}_{k,k-1} = \begin{bmatrix} \hat{a}_{k-1} & \hat{x}_{k-1} \\ 0 & 1 \end{bmatrix} \mathbf{P}_{k-1,k-1} \begin{bmatrix} \hat{a}_{k-1} & 0 \\ \hat{x}_{k-1} & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0.01 \end{bmatrix} \\ \mathbf{G}_k = \mathbf{P}_{k,k-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left[[1 \quad 0] \mathbf{P}_{k,k-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0.01 \right]^{-1} \\ \mathbf{P}_{k,k} = \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \mathbf{G}_k [1 \quad 0] \right] \mathbf{P}_{k,k-1} \\ \begin{bmatrix} \hat{x}_k \\ \hat{a}_k \end{bmatrix} = \begin{bmatrix} \hat{a}_{k-1} & \hat{x}_{k-1} \\ \hat{a}_{k-1} & \hat{a}_{k-1} \end{bmatrix} + \mathbf{G}_k (v_k - \hat{a}_{k-1} \hat{x}_{k-1}) \end{array} \right. \quad (8.17)$$

^① 在本书的翻译过程中,原作者对本节做了少量更正和修改。

式中, x_0 的初始估计是 $\hat{x}_0 = E(x_0) = 1$, 但是 \hat{a}_0 未知。

为了测试该自适应系统辨识算法, 我们制造两个伪随机序列 $\{\eta_k\}$ 和 $\{\zeta_k\}$, 它们的均值为零, 方差为上面的给定值并且假定 a 的值为 -1 用于生成数据 $\{v_k\}$ 。

为了应用(8.17)中描述的算法, 需要 a 的初始估计 \hat{a}_0 。对已生成的数据 $v_1 = -1.1$, 应用式(8.14)。可得

$$0.09a^2 + 2.2a + 1.2 = 0$$

或者

$$a_0 = -1.261, -0.961$$

事实上, 初始估计 \hat{a}_0 可以任意选取。在图 8.3 中, 我们画出了 \hat{a}_0 取十种不同的值时 \hat{a}_k 的曲线(当然也可以取 $-1.261, -0.961$ 等)。可以看出当 $k \geq 15$ 时, 基于任何一个初始条件所产生的估计 \hat{a}_k , 都非常接近真实值 $a = -1$ 。

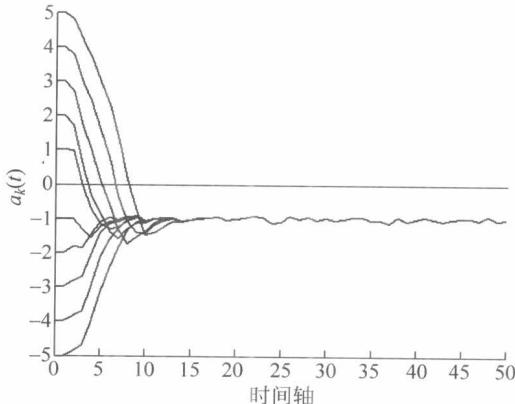


图 8.3 系统辨识

假如 a 是确定的, 从而有 $a_{k+1} = a_k$, 则式(8.16)变为

$$\left\{ \begin{array}{l} \begin{bmatrix} x_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} a_k x_k \\ a_k \end{bmatrix} \\ v_k = [1 \quad 0] \begin{bmatrix} x_k \\ a_k \end{bmatrix} + \eta_k \end{array} \right.$$

则卡尔曼增益矩阵变为

$$\mathbf{G}_k = \begin{bmatrix} g_k \\ 0 \end{bmatrix}$$

g_k 为常值, 并由滤波算法中的“相关”公式, 有

$$\begin{bmatrix} \hat{x}_k \\ \hat{a}_k \end{bmatrix} = \begin{bmatrix} \hat{a}_{k-1} & \hat{x}_{k-1} \\ \hat{a}_{k-1} \end{bmatrix} + \begin{bmatrix} g_k \\ 0 \end{bmatrix} (v_k - \hat{a}_{k-1} \hat{x}_{k-1})$$

注意到, 因为 $\hat{a}_k = \hat{a}_{k-1} = \hat{a}_0$ 不依赖于数据 $\{v_k\}$, 所以不能辨识出 a 的值。

8.5 改进的扩展卡尔曼滤波

在本节将改进前面章节中讨论的扩展卡尔曼滤波算法, 并介绍一个效率更高的并行计算方案, 用于系统参数辨识。改进是通过一个线性化步骤完成的。改进卡尔曼滤波算法能够应用于实时系统参数辨识, 甚至可以用于时变随机系统。给出两个数例并进行计算机仿真, 来证明改进滤波方案比原始扩展卡尔曼滤波算法效率更高。

所考虑的非线性时变系统如下:

$$\left\{ \begin{array}{l} \begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_k(\mathbf{y}_k) \mathbf{x}_k \\ \mathbf{H}_k(\mathbf{x}_k, \mathbf{y}_k) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Gamma}_k^1(\mathbf{x}_k, \mathbf{y}_k) & \mathbf{0} \\ \boldsymbol{\Gamma}_k^2(\mathbf{x}_k, \mathbf{y}_k) & \boldsymbol{\Gamma}_k^3(\mathbf{x}_k, \mathbf{y}_k) \end{bmatrix} \begin{bmatrix} \underline{\xi}_k^1 \\ \underline{\xi}_k^2 \end{bmatrix} \\ \mathbf{v}_k = [\mathbf{C}_k(\mathbf{x}_k, \mathbf{y}_k) \quad \mathbf{0}] \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} + \underline{\eta}_k \end{array} \right. \quad (8.18)$$

式中 \mathbf{x}_k 、 \mathbf{y}_k 分别是 n 维和 m 维向量, $\begin{bmatrix} \underline{\xi}_k^1 \\ \underline{\xi}_k^2 \end{bmatrix}$ 和 $\{\underline{\eta}_k\}$ 是不相关的零均值高斯白噪声序列, 方差阵分别为

$$\mathbf{Q}_k = \text{Var} \begin{bmatrix} \underline{\xi}_k^1 \\ \underline{\xi}_k^2 \end{bmatrix}, \quad \mathbf{R}_k = \text{Var}(\underline{\eta}_k)$$

\mathbf{F}_k 、 \mathbf{H}_k 、 $\boldsymbol{\Gamma}_k^1$ 、 $\boldsymbol{\Gamma}_k^2$ 、 $\boldsymbol{\Gamma}_k^3$ 和 \mathbf{C}_k 分别是非线性矩阵值函数。假设 \mathbf{F}_k 和 \mathbf{C}_k 是可微的。

改进卡尔曼滤波算法包括两个子算法。下面介绍的算法 I 是前面讨论的扩展卡尔曼滤波的一个改进。它和以前算法的不同之处是实时线性泰勒近似不是在前一估计处完成的, 为了提高性能, 泰勒近似是在由标准卡尔曼滤波算法(在下面被称作算法 II)得到的 \mathbf{x}_k 的最优估计处完成的, 它是根据(8.18)的子系统

$$\left\{ \begin{array}{l} \mathbf{x}_{k+1} = \mathbf{F}_k(\bar{\mathbf{y}}_k) \mathbf{x}_k + \boldsymbol{\Gamma}_k^1(\bar{\mathbf{x}}_k, \bar{\mathbf{y}}_k) \underline{\xi}_k^1 \\ \mathbf{v}_k = \mathbf{C}(\bar{\mathbf{x}}_k, \bar{\mathbf{y}}_k) \mathbf{x}_k + \underline{\eta}_k \end{array} \right. \quad (8.19)$$

由算法 I 中的估计值 $(\tilde{x}_k, \tilde{y}_k)$ 计算。换句话说, 这两个算法以相同的初始估计并行运行, 如图 8.4 所示, 其中算法 I (即改进卡尔曼滤波算法) 产生估计 $\begin{bmatrix} \tilde{x}_k \\ \tilde{y}_k \end{bmatrix}$, 其输入 \hat{x}_{k-1} 则由算法 II (即线性系统(8.19)中的标准卡尔曼滤波算法) 给出; 算法 II 用于产生估计 \hat{x}_k , 其输入 $\begin{bmatrix} \tilde{x}_{k-1} \\ \tilde{y}_{k-1} \end{bmatrix}$ 由算法 I 给出。两个算法的合成被称作并行算法(I 和 II)。

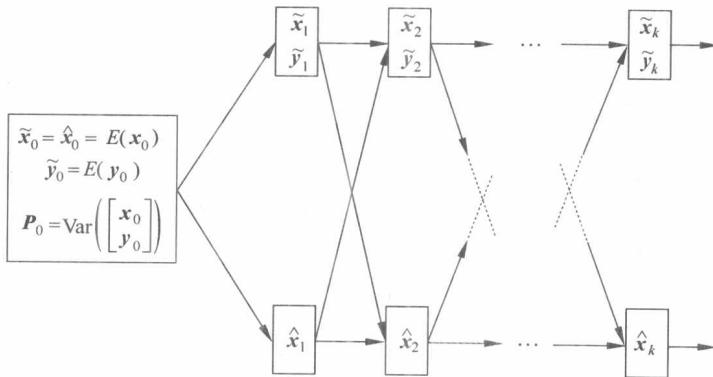


图 8.4 改进的扩展卡尔曼滤波并行算法

算法 I

设置

$$\begin{bmatrix} \tilde{x}_0 \\ \tilde{y}_0 \end{bmatrix} = \begin{bmatrix} E(x_0) \\ E(y_0) \end{bmatrix}, \quad \mathbf{P}_0 = \text{Var}\left(\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}\right)$$

对于 \$k=1, 2, \dots\$, 计算:

$$\begin{aligned} \mathbf{P}_{k,k-1} &= \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{y}_{k-1} \end{bmatrix}} \begin{bmatrix} \mathbf{F}_{k-1}(\tilde{y}_{k-1}) \hat{x}_{k-1} \\ \mathbf{H}_{k-1}(\hat{x}_{k-1}, \tilde{y}_{k-1}) \end{bmatrix} \right] \mathbf{P}_{k-1} \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{y}_{k-1} \end{bmatrix}} \begin{bmatrix} \mathbf{F}_{k-1}(\tilde{y}_{k-1}) \hat{x}_{k-1} \\ \mathbf{H}_{k-1}(\hat{x}_{k-1}, \tilde{y}_{k-1}) \end{bmatrix} \right]^T \\ &+ \begin{bmatrix} \boldsymbol{\Gamma}_{k-1}^1(\tilde{x}_{k-1}, \tilde{y}_{k-1}) & \mathbf{0} \\ \boldsymbol{\Gamma}_{k-1}^2(\tilde{x}_{k-1}, \tilde{y}_{k-1}) & \boldsymbol{\Gamma}_{k-1}^3(\tilde{x}_{k-1}, \tilde{y}_{k-1}) \end{bmatrix} \\ &\cdot \mathbf{Q}_{k-1} \begin{bmatrix} \boldsymbol{\Gamma}_{k-1}^1(\tilde{x}_{k-1}, \tilde{y}_{k-1}) & \mathbf{0} \\ \boldsymbol{\Gamma}_{k-1}^2(\tilde{x}_{k-1}, \tilde{y}_{k-1}) & \boldsymbol{\Gamma}_{k-1}^3(\tilde{x}_{k-1}, \tilde{y}_{k-1}) \end{bmatrix}^T \end{aligned}$$

$$\begin{bmatrix} \tilde{x}_{k|k-1} \\ \tilde{y}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{k-1}(\tilde{y}_{k-1}) \hat{x}_{k-1} \\ \mathbf{H}_{k-1}(\hat{x}_{k-1}, \tilde{y}_{k-1}) \end{bmatrix}$$

$$\begin{aligned}
 \mathbf{G}_k &= \mathbf{P}_{k,k-1} \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix}} \mathbf{C}_k(\tilde{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{y}}_{k|k-1}) \right]^T \left\{ \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix}} \mathbf{C}_k(\tilde{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{y}}_{k|k-1}) \right] \right. \\
 &\quad \cdot \mathbf{P}_{k,k-1} \left. \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix}} \mathbf{C}_k(\tilde{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{y}}_{k|k-1}) \right]^T + \mathbf{R}_k \right\}^{-1} \\
 \mathbf{P}_k &= \left[\mathbf{I} - \mathbf{G}_k \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix}} \mathbf{C}_k(\tilde{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{y}}_{k|k-1}) \right] \right] \mathbf{P}_{k,k-1} \\
 \begin{bmatrix} \tilde{\mathbf{x}}_k \\ \tilde{\mathbf{y}}_k \end{bmatrix} &= \begin{bmatrix} \tilde{\mathbf{x}}_{k|k-1} \\ \tilde{\mathbf{y}}_{k|k-1} \end{bmatrix} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k(\tilde{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{y}}_{k|k-1}) \tilde{\mathbf{x}}_{k|k-1})
 \end{aligned}$$

式中 $\mathbf{Q}_k = \text{Var} \begin{pmatrix} \xi_k^1 \\ \xi_k^2 \end{pmatrix}$, $\mathbf{R}_k = \text{Var}(\eta_k)$, $\hat{\mathbf{x}}_{k-1}$ 由下面的算法 II 得出。

算法 II

设置

$$\hat{\mathbf{x}}_0 = E(\mathbf{x}_0), \quad \mathbf{P}_0 = \text{Var}(\mathbf{x}_0)$$

对于 $k=1, 2, \dots$, 计算:

$$\begin{aligned}
 \mathbf{P}_{k,k-1} &= [\mathbf{F}_{k-1}(\tilde{\mathbf{y}}_{k-1})] \mathbf{P}_{k-1} [\mathbf{F}_{k-1}(\tilde{\mathbf{y}}_{k-1})]^T \\
 &\quad + [\mathbf{T}_{k-1}^l(\tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1})] \mathbf{Q}_{k-1} [\mathbf{T}_{k-1}^l(\tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1})]^T \\
 \hat{\mathbf{x}}_{k|k-1} &= [\mathbf{F}_{k-1}(\tilde{\mathbf{y}}_{k-1})] \hat{\mathbf{x}}_{k-1} \\
 \mathbf{G}_k &= \mathbf{P}_{k,k-1} [\mathbf{C}_k(\tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1})]^T \\
 &\quad \cdot [[\mathbf{C}_k(\tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1})] \mathbf{P}_{k,k-1} [\mathbf{C}_k(\tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1})]^T + \mathbf{R}_k]^{-1} \\
 \mathbf{P}_k &= [\mathbf{I} - \mathbf{G}_k [\mathbf{C}_k(\tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1})]] \mathbf{P}_{k,k-1} \\
 \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - [\mathbf{C}_k(\tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1})] \hat{\mathbf{x}}_{k|k-1})
 \end{aligned}$$

其中 $\mathbf{Q}_k = \text{Var}(\xi_k^1)$, $\mathbf{R}_k = \text{Var}(\eta_k)$, $(\tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1})$ 由算法 I 得到。

这里使用了下面的符号

$$\left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{y}_{k-1} \end{bmatrix}} \begin{bmatrix} \mathbf{F}_{k-1}(\tilde{\mathbf{y}}_{k-1}) \hat{\mathbf{x}}_{k-1} \\ \mathbf{H}_{k-1}(\hat{\mathbf{x}}_{k-1}, \tilde{\mathbf{y}}_{k-1}) \end{bmatrix} \right] = \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{y}_{k-1} \end{bmatrix}} \begin{bmatrix} \mathbf{F}_k(\mathbf{y}_{k-1}) \mathbf{x}_{k-1} \\ \mathbf{H}_{k-1}(\mathbf{x}_{k-1}, \mathbf{y}_{k-1}) \end{bmatrix} \right]_{\substack{\mathbf{x}_{k-1} = \hat{\mathbf{x}}_{k-1} \\ \mathbf{y}_{k-1} = \tilde{\mathbf{y}}_{k-1}}}$$

改进卡尔曼滤波算法(即算法 I)不同于原始卡尔曼滤波算法之处在于: 在每一个时刻,(非线性)向量值函数 $\mathbf{F}_{k-1}(\mathbf{y}_{k-1}) \mathbf{x}_{k-1}$ 的雅克比矩阵和预测值

$\begin{bmatrix} \tilde{\mathbf{x}}_{k|k-1} \\ \tilde{\mathbf{y}}_{k|k-1} \end{bmatrix}$ 都在最优估计 $\hat{\mathbf{x}}_{k-1}$ 处进行估计。其中 $\hat{\mathbf{x}}_{k-1}$ 由标准卡尔曼滤波算法(算法 II)确定。

下面给出改进卡尔曼滤波算法的一个推导。令 $\hat{\mathbf{x}}_k$ 是线性系统(8.19)中状态向量 \mathbf{x}_k 的最优估计。对于使用了所有数据信息 v_1, \dots, v_k 的 \mathbf{x}_k 的所有线性无偏估计 \mathbf{z}_k , 有下式:

$$\text{Var}(\hat{\mathbf{x}}_k - \mathbf{x}_k) \leq \text{Var}(\mathbf{z}_k - \mathbf{x}_k) \quad (8.20)$$

因为(8.19)是原始系统在 $(\bar{\mathbf{x}}_k, \bar{\mathbf{y}}_k)$ 处的子系统, 正如推导扩展卡尔曼滤波一样, 有

$$\text{Var}(\hat{\mathbf{X}}_k - \mathbf{X}_k) \leq \text{Var}(\tilde{\mathbf{X}}_k - \mathbf{X}_k) \quad (8.21)$$

式中

$$\mathbf{X}_k = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix}, \quad \hat{\mathbf{X}}_k = \begin{bmatrix} \hat{\mathbf{x}}_k \\ \bar{\mathbf{y}}_k \end{bmatrix}, \quad \tilde{\mathbf{X}}_k = \begin{bmatrix} \tilde{\mathbf{x}}_k \\ \bar{\mathbf{y}}_k \end{bmatrix}$$

现在, 考虑定义在 \mathbb{R}^{n+m} 上的一个 $(n+m)$ 维非线性可微向量值函数:

$$\mathbf{Z} = f(\mathbf{X}) \quad (8.22)$$

因为目的是从 \mathbf{X}_k 的某个(最优)估计 $\hat{\mathbf{X}}_k$ 中估计:

$$\mathbf{Z}_k = f(\mathbf{X}_k)$$

以 $\hat{\mathbf{X}}_k$ 作为线性泰勒近似的中心点。进行以上操作后, 选择 \mathbf{X}_k 中一个更好的估计作为中心点将会得到 \mathbf{Z}_k 更好的估计值。换句话说, 如果用 $\hat{\mathbf{X}}_k$ 代替 $\tilde{\mathbf{X}}_k$ 作为 \mathbf{Z}_k 线性泰勒近似的中心点, 将会得到 \mathbf{Z}_k 更好的估计值, 如图 8.5 所示。在这里, $\tilde{\mathbf{X}}_k$ 被用来作为标准卡尔曼滤波中线性泰勒近似的中心点。

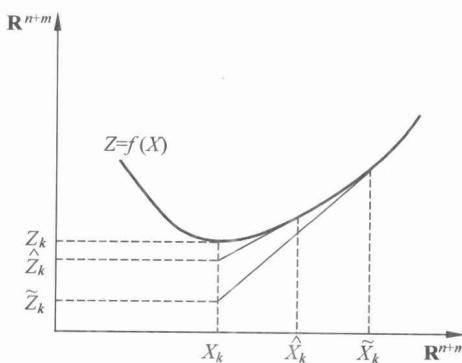


图 8.5 线性化示意图

如果 $f(\mathbf{X})$ 在 \mathbf{X} 处的雅克比矩阵是 $f'(\mathbf{X})$, 那么以 $\hat{\mathbf{X}}_k$ 为重心点, \mathbf{Z}_k 的线性泰勒估计 $\hat{\mathbf{Z}}_k$ 可以由下式给出:

$$\hat{\mathbf{Z}}_k = f(\hat{\mathbf{X}}_k) + f'(\hat{\mathbf{X}}_k)(\mathbf{X}_k - \hat{\mathbf{X}}_k)$$

或者写成

$$\hat{\mathbf{Z}}_k = f'(\hat{\mathbf{X}}_k)\mathbf{X}_k + (f(\hat{\mathbf{X}}_k) - f'(\hat{\mathbf{X}}_k)\hat{\mathbf{X}}_k) \quad (8.23)$$

下面回到非线性模型(8.18), 并将线性公式(8.23)应用到下式定义的 f_k :

$$f_k \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{F}_k(\mathbf{y}_k)\mathbf{x}_k \\ \mathbf{H}_k(\mathbf{x}_k, \mathbf{y}_k) \end{bmatrix} \quad (8.24)$$

假设 $(\bar{\mathbf{x}}_k, \bar{\mathbf{y}}_k)$ 和 $\hat{\mathbf{x}}_k$ 能够在 k 时刻由并行算法(I和II)求出。转到第 $k+1$ 步, 以 $\begin{bmatrix} \hat{\mathbf{x}}_k \\ \bar{\mathbf{y}}_k \end{bmatrix}$ (代替 $\begin{bmatrix} \bar{\mathbf{x}}_k \\ \bar{\mathbf{y}}_k \end{bmatrix}$) 为重心点, 根据式(8.23), 并使用雅克比矩阵对模型(8.18)进行线性化:

$$f'_k \begin{bmatrix} \hat{\mathbf{x}}_k \\ \bar{\mathbf{y}}_k \end{bmatrix} = \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix}} \begin{bmatrix} \mathbf{F}_k(\bar{\mathbf{y}}_k)\hat{\mathbf{x}}_k \\ \mathbf{H}_k(\hat{\mathbf{x}}_k, \bar{\mathbf{y}}_k) \end{bmatrix} \right]$$

此外, 正如在标准扩展卡尔曼滤波中常做的那样, 在模型(8.18)的线性化中, 对于矩阵 $\mathbf{C}_k(\mathbf{x}_k, \mathbf{y}_k)$ 、 $\boldsymbol{\Gamma}_k^i(\mathbf{x}_k, \mathbf{y}_k)$ ($i=1, 2, 3$), 分别做零阶泰勒近似, 得到 $\mathbf{C}_k(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)$ 、 $\boldsymbol{\Gamma}_k^i(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)$ 。可得下面的线性状态空间描述:

$$\begin{cases} \begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_{k+1} \end{bmatrix} = \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix}} \begin{bmatrix} \mathbf{F}_k(\bar{\mathbf{y}}_k)\hat{\mathbf{x}}_k \\ \mathbf{H}_k(\hat{\mathbf{x}}_k, \bar{\mathbf{y}}_k) \end{bmatrix} \right] \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} + \mathbf{u}_k + \begin{bmatrix} \boldsymbol{\Gamma}_k^1(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k) & \mathbf{0} \\ \boldsymbol{\Gamma}_k^2(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k) & \boldsymbol{\Gamma}_k^3(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k) \end{bmatrix} \begin{bmatrix} \xi_k^1 \\ \xi_k^2 \end{bmatrix} \\ \mathbf{v}_k = [\mathbf{C}_k(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k) \quad \mathbf{0}] \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} + \eta_k \end{cases} \quad (8.25)$$

式中, 常值向量

$$\mathbf{u}_k = \begin{bmatrix} \mathbf{F}_k(\bar{\mathbf{y}}_k)\hat{\mathbf{x}}_k \\ \mathbf{H}_k(\hat{\mathbf{x}}_k, \bar{\mathbf{y}}_k) \end{bmatrix} - \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix}} \begin{bmatrix} \mathbf{F}_k(\bar{\mathbf{y}}_k)\hat{\mathbf{x}}_k \\ \mathbf{H}_k(\hat{\mathbf{x}}_k, \bar{\mathbf{y}}_k) \end{bmatrix} \right] \begin{bmatrix} \hat{\mathbf{x}}_k \\ \bar{\mathbf{y}}_k \end{bmatrix}$$

是确定性控制输入。因此, 类似于标准卡尔曼滤波的推导, 得到系统(8.25)的算法I。其中算法I的预测项 $\begin{bmatrix} \tilde{\mathbf{x}}_{k|k-1} \\ \tilde{\mathbf{y}}_{k|k-1} \end{bmatrix}$ 是唯一包含 \mathbf{u}_k 的项, 如下式所示:

$$\begin{aligned}
\begin{bmatrix} \tilde{\mathbf{x}}_{k|k-1} \\ \tilde{\mathbf{y}}_{k|k-1} \end{bmatrix} &= \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{y}_{k-1} \end{bmatrix}} \begin{bmatrix} \mathbf{F}_{k-1}(\bar{\mathbf{y}}_{k-1}) \hat{\mathbf{x}}_{k-1} \\ \mathbf{H}_{k-1}(\hat{\mathbf{x}}_{k-1}, \bar{\mathbf{y}}_{k-1}) \end{bmatrix} \right] \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \bar{\mathbf{y}}_{k-1} \end{bmatrix} + \mathbf{u}_{k-1} \\
&= \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{y}_{k-1} \end{bmatrix}} \begin{bmatrix} \mathbf{F}_{k-1}(\bar{\mathbf{y}}_{k-1}) \hat{\mathbf{x}}_{k-1} \\ \mathbf{H}_{k-1}(\hat{\mathbf{x}}_{k-1}, \bar{\mathbf{y}}_{k-1}) \end{bmatrix} \right] \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \bar{\mathbf{y}}_{k-1} \end{bmatrix} + \begin{bmatrix} \mathbf{F}_{k-1}(\bar{\mathbf{y}}_{k-1}) \hat{\mathbf{x}}_{k-1} \\ \mathbf{H}_{k-1}(\hat{\mathbf{x}}_{k-1}, \bar{\mathbf{y}}_{k-1}) \end{bmatrix} \\
&\quad - \left[\frac{\partial}{\partial \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{y}_{k-1} \end{bmatrix}} \begin{bmatrix} \mathbf{F}_{k-1}(\bar{\mathbf{y}}_{k-1}) \hat{\mathbf{x}}_{k-1} \\ \mathbf{H}_{k-1}(\hat{\mathbf{x}}_{k-1}, \bar{\mathbf{y}}_{k-1}) \end{bmatrix} \right] \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \bar{\mathbf{y}}_{k-1} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{F}_{k-1}(\bar{\mathbf{y}}_{k-1}) \hat{\mathbf{x}}_{k-1} \\ \mathbf{H}_{k-1}(\hat{\mathbf{x}}_{k-1}, \bar{\mathbf{y}}_{k-1}) \end{bmatrix}
\end{aligned}$$

8.6 时变参数辨识

在本节中,给出两个数值例子,来说明改进的扩展卡尔曼滤波相比于标准卡尔曼滤波在状态估计和系统参数辨识方面的优势。

给出一个非线性系统

$$\begin{cases} \begin{bmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & z_k \\ -0.1 & 1 \\ & z_k \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} + \begin{bmatrix} \xi^1_k \\ \xi^2_k \\ \xi^3_k \end{bmatrix}, & \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.0 \\ 0.1 \end{bmatrix} \\ v_k = x_k + \eta_k \end{cases}$$

式中 ξ_k 和 η_k 是不相关的零均值高斯白噪声序列,且对所有 k ,有 $\text{Var}(\xi_k) = 0.1 \mathbf{I}_3$ 和 $\text{Var}(\eta_k) = 0.01$ 。

产生两个伪随机噪声序列,并用

$$\begin{bmatrix} \hat{x}_0 \\ \hat{y}_0 \\ \hat{z}_0 \end{bmatrix} = \begin{bmatrix} \tilde{x}_0 \\ \tilde{y}_0 \\ \tilde{z}_0 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \\ 1.0 \end{bmatrix} \quad \text{和} \quad \mathbf{P}_0 = \mathbf{I}_3$$

作为比较扩展卡尔曼滤波算法和并行算法(I和II)时的初值。

计算机仿真结果表明,两种方法都能够精确估计状态向量的实际分量 x_k 。

但是,比起标准的扩展卡尔曼滤波算法,并行算法(I和II)能更好地估计 y_k 和 z_k 分量。

在图8.6和图8.7中,分别用标准的扩展卡尔曼滤波算法和并行算法画出 y_k 的估计值并与 y_k 的实际值做比较。同样,用两种方法分别画出 z_k 部分的估计值,如图8.8和图8.9所示。

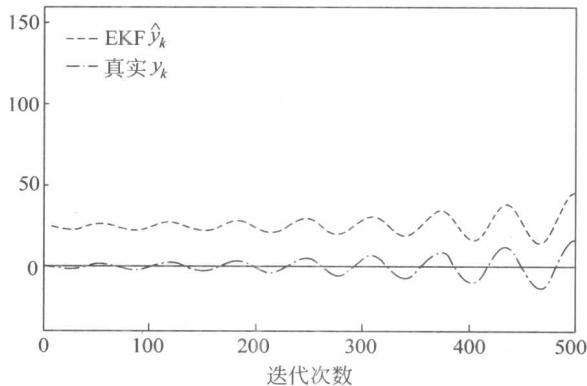


图8.6 扩展卡尔曼滤波对 y_k 的辨识结果

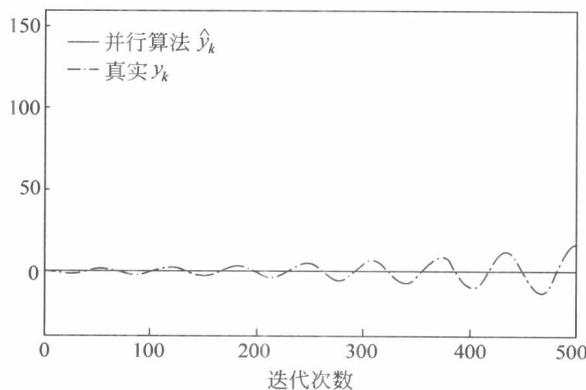
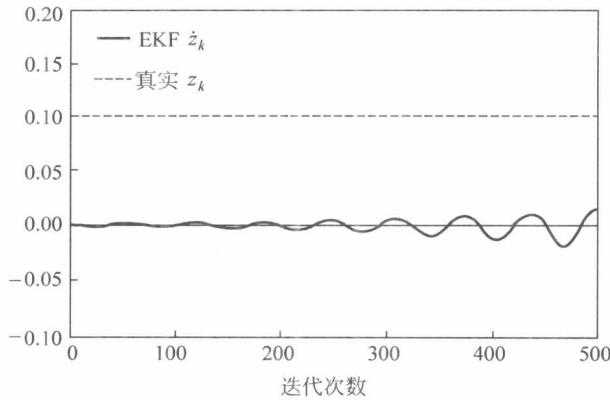
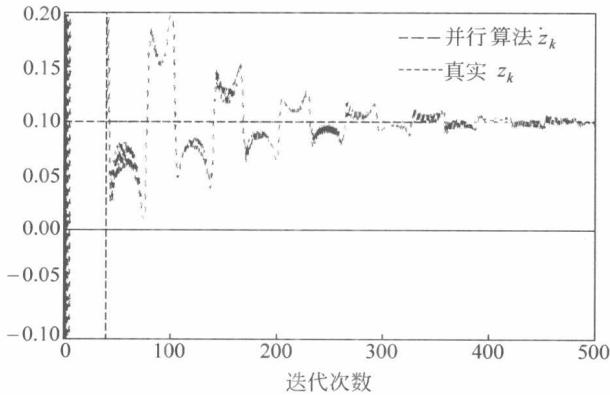


图8.7 并行算法对 y_k 的辨识结果

图 8.8 扩展卡尔曼滤波对 z_k 的辨识结果图 8.9 并行算法对 z_k 的辨识结果

第二个例子,考虑以下系统参数识别问题,时变随机系统(8.19)含有一个未知的时变参数向量 $\underline{\theta}_k$ 。通常,只假设 $\underline{\theta}_k$ 是一个零均值随机向量。用 $\underline{\theta}_k$ 取代非线性模型(8.25)中的 y_k 。对该模型应用改进的扩展卡尔曼滤波算法,可得

算法 I' 令 $\text{Var}(\underline{\theta}_0) > \mathbf{0}$,且设置

$$\begin{bmatrix} \hat{x}_0 \\ \hat{\theta}_0 \end{bmatrix} = \begin{bmatrix} E(x_0) \\ \mathbf{0} \end{bmatrix}, \quad P_{0,0} = \begin{bmatrix} \text{Var}(x_0) & \mathbf{0} \\ \mathbf{0} & \text{Var}(\underline{\theta}_0) \end{bmatrix}$$

当 $k=1,2,\dots$ 时,计算:

$$\begin{bmatrix} \hat{x}_{k|k-1} \\ \hat{\theta}_{k|k-1} \end{bmatrix} = \begin{bmatrix} A_{k-1}(\hat{\theta}_{k-1}) \hat{x}_{k-1} \\ \hat{\theta}_{k-1} \end{bmatrix}$$

$$\begin{aligned}
 \mathbf{P}_{k,k-1} &= \left[\begin{array}{cc} \mathbf{A}_{k-1}(\tilde{\boldsymbol{\theta}}_{k-1}) & \frac{\partial}{\partial \boldsymbol{\theta}} [\mathbf{A}_{k-1}(\tilde{\boldsymbol{\theta}}_{k-1}) \hat{\mathbf{x}}_{k-1}] \\ \mathbf{0} & \mathbf{I} \end{array} \right] \mathbf{P}_{k-1,k-1} \\
 &\cdot \left[\begin{array}{cc} \mathbf{A}_{k-1}(\tilde{\boldsymbol{\theta}}_{k-1}) & \frac{\partial}{\partial \boldsymbol{\theta}} [\mathbf{A}_{k-1}(\tilde{\boldsymbol{\theta}}_{k-1}) \hat{\mathbf{x}}_{k-1}] \\ \mathbf{0} & \mathbf{I} \end{array} \right]^T \\
 &+ \left[\begin{array}{cc} \mathbf{\Gamma}_{k-1}(\tilde{\boldsymbol{\theta}}_{k-1}) \mathbf{Q}_{k-1} \mathbf{\Gamma}_{k-1}^T(\tilde{\boldsymbol{\theta}}_{k-1}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array} \right]
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{G}_k &= \mathbf{P}_{k,k-1} [\mathbf{C}_k(\tilde{\boldsymbol{\theta}}_{k|k-1}) \quad \mathbf{0}]^T \\
 &\cdot [[\mathbf{C}_k(\tilde{\boldsymbol{\theta}}_{k|k-1}) \quad \mathbf{0}] \mathbf{P}_{k,k-1} [\mathbf{C}_k(\tilde{\boldsymbol{\theta}}_{k|k-1}) \quad \mathbf{0}]^T + \mathbf{R}_k]^{-1} \\
 \mathbf{P}_{k,k} &= [\mathbf{I} - \mathbf{G}_k [\mathbf{C}_k(\tilde{\boldsymbol{\theta}}_{k|k-1}) \quad \mathbf{0}]] \mathbf{P}_{k,k-1} \\
 \begin{bmatrix} \tilde{\mathbf{x}}_k \\ \tilde{\boldsymbol{\theta}}_k \end{bmatrix} &= \begin{bmatrix} \tilde{\mathbf{x}}_{k|k-1} \\ \tilde{\boldsymbol{\theta}}_{k|k-1} \end{bmatrix} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k(\tilde{\boldsymbol{\theta}}_{k|k-1}) \tilde{\mathbf{x}}_{k|k-1})
 \end{aligned}$$

式中, $\hat{\mathbf{x}}_k$ 是在并行算法 II 中用 $\tilde{\boldsymbol{\theta}}_k$ 代替 $\bar{\mathbf{y}}_k$ 得到的。

为了检验并行算法(I'和II)的性能, 考虑下面含有未知系统参数 θ_k 的随机系统:

$$\begin{cases} \mathbf{x}_{k+1} = \begin{bmatrix} 1 & 1 \\ 0 & \theta_k \end{bmatrix} \mathbf{x}_k + \xi_k, & \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \mathbf{v}_k = [1 \quad 0] \mathbf{x}_k + \eta_k \end{cases}$$

计算机仿真时, 使用由下式给出的变值参数 θ_k

$$\theta_k = \begin{cases} 1.0, & 0 \leq k \leq 20 \\ 1.2, & 20 < k \leq 40 \\ 1.5, & 40 < k \leq 60 \quad \text{及} \quad \hat{\theta}_0 = 5.0 \\ 2.0, & 60 < k \leq 80 \\ 2.3, & 80 < k \end{cases}$$

现在的问题是辨识这些未知量。事实证明标准的扩展卡尔曼滤波算法不能合理估计 θ_k , 且估计值可能随着 k 的增大而成指数增长。在图 8.10 中, θ_k 的估计值是应用并行算法(I'和II)得到的, 其初始差值为(i) $\text{Var}(\theta_0) = 0.1$; (ii) $\text{Var}(\theta_0) = 1.0$; (iii) $\text{Var}(\theta_0) = 50$ 。

最后指出, 所有现有的扩展卡尔曼滤波算法设计都是有针对性的, 这是由于

推导结果使用了不同的线性化方法引起的。因此，通常没有严格的理论能保证扩展或改进的扩展卡尔曼滤波算法的最优性。

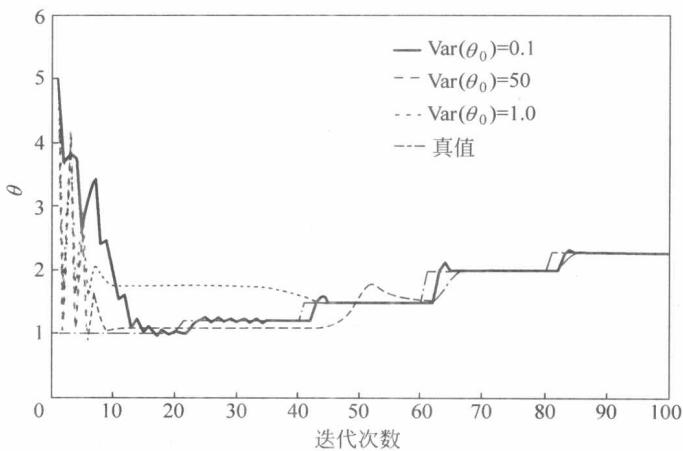


图 8.10 时变参数辨识结果

练习

8.1 考虑图 8.11 所示的两维雷达跟踪系统。为了简便，假设导弹在 y 轴的正向运动，则 $\dot{x}=0$, $\dot{y}=v$ 和 $\ddot{y}=a$, 其中 v 和 a 分别表示导弹的速度和加速度。

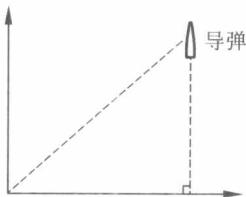


图 8.11 雷达跟踪系统

(a) 假设雷达安装在原点，测量值为距离 r 和角位移 θ , (r, θ) 为极坐标表示。推导这个雷达跟踪模型的非线性方程：

$$\begin{cases} \dot{r} = v \sin \theta \\ \dot{\theta} = \frac{v}{r} \cos \theta \end{cases}, \quad \begin{cases} \ddot{r} = a \sin \theta + \frac{v^2}{r^2} \cos^2 \theta \\ \ddot{\theta} = \left(\frac{ar - v^2 \sin \theta}{r^2} \right) \cos \theta - \frac{v^2}{r^2} \sin \theta \cos \theta \end{cases}$$

(b) 给定状态向量 $x := \begin{bmatrix} r \\ \dot{r} \\ \theta \\ \dot{\theta} \end{bmatrix}$, 建立一个该模型的基于向量的非线性微分方程;

- (c) 假设只有距离 r 能够观测, 系统和观测方程都分别受到随机干扰 $\{\xi\}$ 和 $\{\eta\}$ 。分别用 x_k 、 $(x_{k+1} - x_k)h^{-1}$ 、 ξ_k 和 η_k 替换 x 、 \dot{x} 、 ξ 和 η , 其中 $h > 0$ 表示采样时间, 建立上面模型的离散时间非线性系统;
- (d) 假设上面系统中的 $\{\xi_k\}$ 、 $\{\eta_k\}$ 是零均值不相关的高斯白噪声序列。描述该非线性系统的扩展卡尔曼滤波算法。

8.2 用式(8.5)和式(8.7)定义的矩阵和向量, 验证非线性模型(8.1)可以用线性模型(8.3)近似。

8.3 验证通过对线性模型(8.3)应用标准卡尔曼方程(2.17)或方程(3.25), 可以得到非线性模型(8.1)的扩展卡尔曼滤波算法, 如式(8.8)所示。

8.4 验证方程(8.13)。

8.5 验证方程(8.14)。

8.6 验证(8.15)给出的算法。

8.7 证明: 如果系统(8.10)中的未知常值向量 $\underline{\theta}$ 是确定性的话(即, 对所有 k , 有 $\underline{\theta}_{k+1} = \underline{\theta}_k$), 则算法(8.15)将不能辨识出 $\underline{\theta}$ 。

8.8 考虑一维模型: $\begin{cases} x_{k+1} = x_k + \xi_k \\ v_k = cx_k + \eta_k \end{cases}$, 式中 $E(x_0) = x^0$, $\text{Var}(x_0) = p_0$, $\{\xi_k\}$ 和 $\{\eta_k\}$ 都是零均值高斯白噪声序列, 满足:

$$E(\xi_k \xi_l) = q_k \delta_{kl}, \quad E(\eta_k \eta_l) = r_k \delta_{kl}$$

$$E(\xi_k \eta_l) = E(\xi_k x_0) = E(\eta_k x_0) = 0$$

假设未知常数 c 是一个随机常数, 即: $c_{k+1} = c_k + \zeta_k$, 式中 ζ_k 也是零均值高斯白噪声序列, 且已知方差 $\text{Var}(\zeta_k) = s_k$ 。

给出估计 c 的算法, 并讨论特殊情况 $s_k = s > 0$ 。

第 9 章

滤波方程解耦

极限(或稳态)卡尔曼滤波器为实时估算线性时不变系统的状态向量提供了一个非常有效的方法。如果状态向量具有非常高的维数 n ,但是只对很少的一部分感兴趣时,这种滤波器将提供大量无用的信息。如果能去除这些,可以提高滤波过程的效率。为此本章介绍了一种解耦方法。它可以将一个 n 维的极限卡尔曼滤波算法分解为 n 个独立的一维递推公式,从而可以去掉不关心的项。

9.1 解耦公式

考虑一个线性时不变随机系统:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \boldsymbol{\Gamma}\underline{\xi}_k \\ \mathbf{v}_k = \mathbf{C}\mathbf{x}_k + \underline{\eta}_k \end{cases} \quad (9.1)$$

所有项的定义与第 6 章相同(见 6.1 节)。则极限卡尔曼增益矩阵为

$$\mathbf{G} = \mathbf{P}\mathbf{C}^T(\mathbf{C}\mathbf{P}\mathbf{C}^T + \mathbf{R})^{-1} \quad (9.2)$$

式中 \mathbf{P} 是矩阵 Riccati 方程

$$\mathbf{P} = \mathbf{A}[\mathbf{P} - \mathbf{P}\mathbf{C}^T(\mathbf{C}\mathbf{P}\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\mathbf{P}]\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T \quad (9.3)$$

的正定解,且 \mathbf{x}_k 的稳态估计 $\vec{\mathbf{x}}_k$ 为

$$\begin{aligned} \vec{\mathbf{x}}_k &= \mathbf{A}\vec{\mathbf{x}}_{k-1} + \mathbf{G}(\mathbf{v}_k - \mathbf{C}\mathbf{A}\vec{\mathbf{x}}_{k-1}) \\ &= (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A}\vec{\mathbf{x}}_{k-1} + \mathbf{G}\mathbf{v}_k \end{aligned} \quad (9.4)$$

(见 6.4 节)。式中, \mathbf{Q}, \mathbf{R} 均为方差阵,且由下式定义:

$$\mathbf{Q} = E(\underline{\xi}_k \underline{\xi}_k^T), \quad \mathbf{R} = E(\underline{\eta}_k \underline{\eta}_k^T)$$

对所有 k 都适用(见 6.1 节)。

令 $\Phi = (\mathbf{I} - \mathbf{G}\mathbf{C})\mathbf{A} := [\phi_{ij}]_{n \times n}, \mathbf{G} = [g_{ij}]_{n \times q}, 1 \leq q \leq n$, 设

$$\vec{x}_k = \begin{bmatrix} x_{k,1} \\ \vdots \\ x_{k,n} \end{bmatrix}, \quad \mathbf{v}_k = \begin{bmatrix} v_{k,1} \\ \vdots \\ v_{k,q} \end{bmatrix}$$

现在考虑 z -变换:

$$\begin{cases} \mathbf{X}_j = \mathbf{X}_j(z) = \sum_{k=0}^{\infty} x_{k,j} z^{-k}, & j = 1, 2, \dots, n \\ \mathbf{V}_j = \mathbf{V}_j(z) = \sum_{k=0}^{\infty} v_{k,j} z^{-k}, & j = 1, 2, \dots, n \end{cases} \quad (9.5)$$

分别对应于 $\{\vec{x}_k\}$ 和 $\{\mathbf{v}_k\}$ 的第 j 个分量。

对于 $k=0, 1, \dots$, 式(9.4)可以改写为

$$x_{k+1,j} = \sum_{i=1}^n \phi_{ji} x_{k,i} + \sum_{i=1}^q g_{ji} v_{k+1,i}$$

由上式可得

$$z\mathbf{X}_j = \sum_{i=1}^n \phi_{ji} \mathbf{X}_i + z \sum_{i=1}^q g_{ji} \mathbf{V}_i$$

令

$$\Lambda = \Lambda(z) = (z\mathbf{I} - \Phi)$$

可得

$$\Lambda \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} = z\mathbf{G} \begin{bmatrix} \mathbf{V}_1 \\ \vdots \\ \mathbf{V}_q \end{bmatrix} \quad (9.6)$$

对于大的 $|z|$ 值, Λ 是对角线占优矩阵, 因此也是可逆的。克莱姆(Cramer)法则可以用来求解式(9.6)中的 $\mathbf{X}_1, \dots, \mathbf{X}_n$ 。设 Δ_i 为把 Λ 的第 i 列用下式代替所得到的结果:

$$z\mathbf{G} \begin{bmatrix} \mathbf{V}_1 \\ \vdots \\ \mathbf{V}_q \end{bmatrix}$$

$\det \Lambda$ 和 $\det \Delta_i$ 都是 z 的 n 阶多项式, 且

$$(\det \Lambda) \mathbf{X}_i(z) = \det \Delta_i \quad (9.7)$$

式中, $i=1, \dots, n$ 。此外, 还可以写出

$$\det \Lambda = z^n + b_1 z^{n-1} + b_2 z^{n-2} + \dots + b_{n-1} z + b_n \quad (9.8)$$

式中

$$\begin{aligned} b_1 &= -(\lambda_1 + \lambda_2 + \cdots + \lambda_n) \\ b_2 &= (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_1\lambda_n + \lambda_2\lambda_3 + \cdots + \lambda_{n-1}\lambda_n) \\ &\vdots \\ b_n &= (-1)^n \lambda_1\lambda_2\cdots\lambda_n \end{aligned}$$

这里, λ_i ($i=1, 2, \dots, n$) 是矩阵 Φ 的特征值。

同理可得

$$\det \mathbf{A}_i = \left[\sum_{l=0}^n c_l^i z^{n-l} \right] \mathbf{V}_1 + \cdots + \left[\sum_{l=0}^n c_l^i z^{n-l} \right] \mathbf{V}_l \quad (9.9)$$

式中 c_l^i ($l=0, 1, \dots, n, i=1, 2, \dots, q$) 也可以精确地计算出来。

现在, 将式(9.8)和式(9.9)代入式(9.7), 并对两边取 z -反变换, 得到如下递推(去耦)公式:

$$\begin{aligned} x_{k,i} &= -b_1 x_{k-1,i} - b_2 x_{k-2,i} - \cdots - b_n x_{k-n,i} \\ &\quad + c_0^i v_{k,1} + c_1^i v_{k-1,1} + \cdots + c_n^i v_{k-n,1} \\ &\quad \cdots \\ &\quad + c_0^q v_{k,q} + c_1^q v_{k-1,q} + \cdots + c_n^q v_{k-n,q} \end{aligned} \quad (9.10)$$

$i=1, 2, \dots, n$ 。

在滤波过程之前, 可以计算出系数 b_1, b_2, \dots, b_n 和 $c_0^i, c_1^i, \dots, c_n^i, i=1, 2, \dots, q$ 。而在式(9.10)中, 每个 $x_{k,i}$ 只取决于之前的状态变量 $x_{k-1,i}, x_{k-2,i}, \dots, x_{k-n,i}$ 和数据信息, 而与任何其他的状态 $x_{k-l,j}$ 无关 ($j \neq i$)。这意味着滤波公式(9.4)已分解成 n 个一维递推公式。

9.2 实时跟踪

现以 3.5 节研究的实时跟踪为例, 详细说明解耦技术。与 3.5 节和练习 3.8 一致, 该实时跟踪模型可以简化为

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{Ax}_k + \boldsymbol{\xi}_k \\ v_k = \mathbf{Cx}_k + \eta_k \end{cases} \quad (9.11)$$

式中

$$\mathbf{A} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = [1 \ 0 \ 0], \quad h > 0$$

$\{\xi_k\}$ 和 $\{\eta_k\}$ 分别是零均值高斯白噪声序列,且满足

$$E(\xi_k \xi^T) = \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix} \delta_{kl}, \quad E(\eta_k \eta_l) = \sigma_m \delta_{kl}$$

$$E(\xi_k \eta_l) = \mathbf{0}, \quad E(\xi_k x_0^T) = \mathbf{0}, \quad E(\eta_k x_0) = \mathbf{0}$$

式中 $\sigma_p, \sigma_v, \sigma_a \geq 0, \sigma_p + \sigma_v + \sigma_a > 0, \sigma_m > 0$ 。

在第6章也介绍了该系统的极限卡尔曼滤波:

$$\begin{cases} \vec{x}_k = \Phi \vec{x}_{k-1} + v_k G \\ \vec{x}_0 = E(x_0) \end{cases}$$

式中, $\Phi = (I - GC)A$, 且

$$G = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = PC / (C^T PC + \sigma_m) = \frac{1}{P[1,1] + \sigma_m} \begin{bmatrix} P[1,1] \\ P[2,1] \\ P[3,1] \end{bmatrix}$$

$P = [P[i,j]]_{3 \times 3}$ 是下面矩阵 Riccati 方程的正定解:

$$P = A[P - PC(C^T PC + \sigma_m)^{-1} C^T P]A^T + \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix}$$

或

$$P = A \left[P - \frac{1}{P[1,1] + \sigma_m} P \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P \right] A^T + \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix} \quad (9.12)$$

又

$$\Phi = (I - GC)A = \begin{bmatrix} 1 - g_1 & (1 - g_1)h & (1 - g_1)h^2/2 \\ -g_2 & 1 - g_2h & h - g_2h^2/2 \\ -g_3 & -g_3h & 1 - g_3h^2/2 \end{bmatrix}$$

则式(9.6)变为

$$\begin{bmatrix} z - 1 + g_1 & -h + hg_1 & -h^2/2 + h^2g_1/2 \\ g_2 & z - 1 + hg_2 & -h + h^2g_2/2 \\ g_3 & hg_3 & z - 1 + h^2g_3/2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = z \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} V$$

根据克莱姆法则,有

$$X_i = H_i V$$

$i=1,2,3$, 式中

$$\begin{aligned}
 H_1 &= \{g_1 + (g_3 h^2/2 + g_2 h - 2g_1)z^{-1} + (g_3 h^2/2 - g_2 h + g_1)z^{-2}\} \\
 &\quad \cdot \{1 + ((g_1 - 3) + g_2 h + g_3 h^2/2)z^{-1} \\
 &\quad + ((3 - 2g_1) - g_2 h + g_3 h^2/2)z^{-2} + (g_1 - 1)z^{-3}\} \\
 H_2 &= \{g_2 + (hg_3 - 2g_2)z^{-1} + (g_2 - hg_3)z^{-2}\} \\
 &\quad \cdot \{1 + ((g_1 - 3) + g_2 h + g_3 h^2/2)z^{-1} \\
 &\quad + ((3 - 2g_1) - g_2 h + g_3 h^2/2)z^{-2} + (g_1 - 1)z^{-3}\} \\
 H_3 &= \{g_3 - 2g_3 z^{-1} + g_3 z^{-2}\} \\
 &\quad \cdot \{1 + ((g_1 - 3) + g_2 h + g_3 h^2)z^{-1} \\
 &\quad + ((3 - 2g_1) - g_2 h + g_3 h^2/2)z^{-2} + (g_1 - 1)z^{-3}\}
 \end{aligned}$$

因此,设

$$\vec{x}_k = \begin{bmatrix} x_k \\ \dot{x}_k \\ \ddot{x}_k \end{bmatrix}$$

使用 z -反变换,有

$$\begin{aligned}
 x_k &= -((g_1 - 3) + g_2 h + g_3 h^2/2)x_{k-1} - ((3 - 2g_1) - g_2 h + g_3 h^2/2)x_{k-2} \\
 &\quad - (g_1 - 1)x_{k-3} + g_1 v_k + (g_3 h^2/2 + g_2 h - 2g_1)v_{k-1} \\
 &\quad + (g_3 h^2/2 - g_2 h + g_1)v_{k-2} \\
 \dot{x}_k &= -((g_1 - 3) + g_2 h + g_3 h^2/2)\dot{x}_{k-1} - ((3 - 2g_1) - g_2 h + g_3 h^2/2)\dot{x}_{k-2} \\
 &\quad - (g_1 - 1)\dot{x}_{k-3} + g_2 v_k + (hg_3 - 2g_2)v_{k-1} + (g_2 - hg_3)v_{k-2} \\
 \ddot{x}_k &= -((g_1 - 3) + g_2 h + g_3 h^2/2)\ddot{x}_{k-1} - ((3 - 2g_1) - g_2 h + g_3 h^2/2)\ddot{x}_{k-2} \\
 &\quad - (g_1 - 1)\ddot{x}_{k-3} + g_3 v_k - 2g_3 v_{k-1} + g_3 v_{k-2}
 \end{aligned}$$

$k=0,1,\dots$,初始条件为 $x_{-1}, \dot{x}_{-1}, \ddot{x}_{-1}$,式中 $k < 0$ 时 $v_k = 0$, $k < -1$ 时 $x_k = \dot{x}_k = \ddot{x}_k = 0$ (见练习 9.2)。

9.3 α - β - γ 跟踪器

α - β - γ 跟踪器是最常见的跟踪器之一,这是一个“次优”滤波器,可以表示为

$$\begin{cases} \dot{\vec{x}}_k = \mathbf{A} \vec{x}_{k-1} + \mathbf{H}(v_k - \mathbf{C} \mathbf{A} \vec{x}_{k-1}) \\ \vec{x}_0 = E(\mathbf{x}_0) \end{cases} \quad (9.13)$$

式中, $\mathbf{H} = [\alpha \quad \beta/h \quad \gamma/h^2]^T$ (α, β, γ 为常数, 见图 9.1)。在实际应用中, α, β, γ 根据物理模型来选取, 并依赖于使用者的经验。

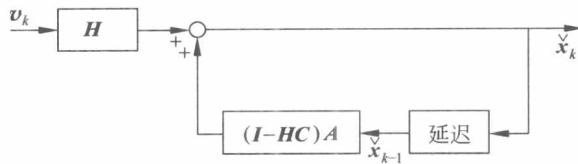


图 9.1 $\alpha\text{-}\beta\text{-}\gamma$ 跟踪器

在这里, 考虑下面的例子

$$\mathbf{A} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \quad \text{和} \quad \mathbf{C} = [1 \quad 0 \quad 0]$$

设

$$g_1 = \alpha, \quad g_2 = \beta/h, \quad g_3 = \gamma/h^2$$

根据 9.2 节推导的解耦滤波公式, 可以得到一个解耦 $\alpha\text{-}\beta\text{-}\gamma$ 跟踪器。可以看到在一定的 α, β 和 γ 值的条件下, 时不变系统(9.11)的 $\alpha\text{-}\beta\text{-}\gamma$ 跟踪器实质上是一个极限卡尔曼滤波器, 由这些条件可以保证跟踪器的“近似最优”性。

由于式(9.12)中的矩阵 \mathbf{P} 是对称的, 可以写为

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{21} & p_{22} & p_{32} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

则式(9.12)变为

$$\mathbf{P} = \mathbf{A} \left[\mathbf{P} - \frac{1}{p_{11} + \sigma_m} \mathbf{P} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{P}^T + \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix} \right] \mathbf{A}^T \quad (9.14)$$

$$\mathbf{G} = \mathbf{P}\mathbf{C}/(\mathbf{C}^T\mathbf{P}\mathbf{C} + \sigma_m) = \frac{1}{p_{11} + \sigma_m} \begin{bmatrix} p_{11} \\ p_{21} \\ p_{31} \end{bmatrix} \quad (9.15)$$

$\alpha\text{-}\beta\text{-}\gamma$ 跟踪器(9.13)成为一个极限卡尔曼滤波器的必要条件是 $\mathbf{H} = \mathbf{G}$, 等价于

$$\begin{bmatrix} \alpha \\ \beta/h \\ \gamma/h^2 \end{bmatrix} = \frac{1}{p_{11} + \sigma_m} \begin{bmatrix} p_{11} \\ p_{21} \\ p_{31} \end{bmatrix}$$

从而

$$\begin{bmatrix} p_{11} \\ p_{21} \\ p_{31} \end{bmatrix} = \frac{\sigma_m}{1-\alpha} \begin{bmatrix} \alpha \\ \beta/h \\ \gamma/h^2 \end{bmatrix} \quad (9.16)$$

另一方面,通过简单的代数运算,由式(9.14)、式(9.15)和式(9.16),可得

$$\begin{aligned} & \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{21} & p_{22} & p_{32} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \\ &= \begin{bmatrix} p_{11} + 2hp_{21} + h^2 p_{31} & p_{21} + hp_{31} + hp_{22} & p_{31} + hp_{32} \\ + h^2 p_{22} + h^3 p_{32} + h^4 p_{33}/4 & + 3h^2 p_{32}/2 + h^3 p_{33}/2 & + h^2 p_{33}/2 \\ p_{21} + hp_{31} + hp_{22} & p_{22} + 2hp_{32} + h^2 p_{33} & p_{32} + hp_{33} \\ + 3h^2 p_{32}/2 + h^3 p_{33}/2 & & \end{bmatrix} \\ & \quad \begin{bmatrix} p_{31} + hp_{32} + h^2 p_{33}/2 & p_{32} + hp_{33} & p_{33} \\ (\alpha + \beta)^2 & (\alpha\beta + \alpha\gamma + \beta^2 & \gamma(\alpha + \beta + \gamma/2)/h^2 \\ + \gamma(\alpha + \beta + \gamma/4) & + 3\beta\gamma/2 + \gamma^2/2)/h & \\ \end{bmatrix} \\ & - \frac{1}{p_{11} + \sigma_m} \begin{bmatrix} (\alpha\beta + \alpha\gamma + \beta^2 & (\beta + \gamma)^2/h^2 & \gamma(\beta + \gamma)/h^3 \\ + 3\beta\gamma/2 + \gamma^2/2)/h & & \\ \gamma(\alpha + \beta + \gamma/2)/h^2 & \gamma(\beta + \gamma)/h^3 & \gamma^2/h^4 \end{bmatrix} \\ &+ \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix} \end{aligned}$$

将式(9.16)代入上面的方程,得到

$$\begin{cases} \frac{h^4}{\gamma^2} \sigma_a = p_{11} + \sigma_m \\ p_{11} = \frac{h^4 (\beta + \gamma)^2}{\gamma^3 (2\alpha + 2\beta + \gamma)} \sigma_a - \frac{h}{\gamma (2\alpha + 2\beta + \gamma)} \sigma_v \\ 2hp_{21} + h^2 p_{31} + h^2 p_{22} = \frac{h^4}{4\gamma^2} (4\alpha^2 + 8\alpha\beta + 2\beta^2 + 4\alpha\gamma + \beta\gamma) \sigma_a + \frac{h^2}{2} \sigma_v - \sigma_p \\ p_{31} + p_{22} = \frac{h^4}{4\gamma^2} (4\alpha + \beta)(\beta + \gamma) \sigma_a + \frac{3}{4} \sigma_v \end{cases} \quad (9.17)$$

及

$$\begin{cases} p_{22} = \frac{\sigma_m}{(1-\alpha)h^2} [\beta(\alpha + \beta + \gamma/4) - \gamma(2 + \alpha)/2] \\ p_{32} = \frac{\sigma_m}{(1-\alpha)h^3} \gamma(\alpha + \beta/2) \\ p_{33} = \frac{\sigma_m}{(1-\alpha)h^4} \gamma(\beta + \gamma) \end{cases} \quad (9.18)$$

联立式(9.16)、式(9.17)和式(9.18),可得

$$\begin{cases} \frac{\sigma_p}{\sigma_m} = \frac{1}{1-\alpha} (\alpha^2 + \alpha\beta + \alpha\gamma/2 - 2\beta) \\ \frac{\sigma_v}{\sigma_m} = \frac{1}{1-\alpha} (\beta^2 - 2\alpha\gamma) h^{-2} \\ \frac{\sigma_a}{\sigma_m} = \frac{1}{1-\alpha} \gamma^2 h^{-4} \end{cases} \quad (9.19)$$

及

$$\mathbf{P} = \frac{\sigma_m}{1-\alpha} \begin{bmatrix} \alpha & \beta/h & \gamma/h^2 \\ \beta/h & (\beta(\alpha + \beta + \gamma/4) - \gamma(2 + \alpha)/2)/h^2 & \gamma(\alpha + \beta/2)/h^3 \\ \gamma/h^2 & \gamma(\alpha + \beta/2)/h^3 & \gamma(\beta + \gamma)/h^4 \end{bmatrix} \quad (9.20)$$

(见练习9.4)。因为 \mathbf{P} 必须是正定的(见定理6.1), α, β 和 γ 的值由下面的方法计算(见练习9.5)。

定理9.1 设 α, β 和 γ 的值满足式(9.19)的条件,并假设 $\sigma_m > 0$ 。则当且仅当下面的条件满足时, α - β - γ 跟踪器是极限卡尔曼滤波器:

(i) $0 < \alpha < 1, \gamma > 0$;

(ii) $\sqrt{2\alpha\gamma} \leq \beta \leq \frac{\alpha}{2-\alpha}(\alpha + \gamma/2)$;

(iii) 下面的矩阵非负定:

$$\tilde{\mathbf{P}} = \begin{bmatrix} \alpha & \beta & \gamma \\ \beta & \beta(\alpha + \beta + \gamma/4) - \gamma(2 + \gamma)/2 & \gamma(\alpha + \beta/2) \\ \gamma & \gamma(\alpha + \beta/2) & \gamma(\beta + \gamma) \end{bmatrix}$$

9.4 一个例子

现在考虑实时跟踪系统(9.11)中 $\sigma_p = \sigma_v = 0, \sigma_a, \sigma_m > 0$ 的特殊情形。综合运用式(9.16)~(9.18),可得

$$\begin{cases} \alpha = 1 - s\gamma^2 \\ \beta^2 = 2\alpha\gamma \\ \alpha^2 + \alpha\beta + \alpha\gamma/2 - 2\beta = 0 \end{cases} \quad (9.21)$$

式中

$$s = \frac{\sigma_m}{\sigma_a} h^{-4}$$

(见练习 9.6)。对式(9.21)做简单的代数运算,可得

$$f(\gamma) := s^3\gamma^6 + s^2\gamma^5 - 3s(s-1/12)\gamma^4 + 6s\gamma^3 + 3(s-1/12)\gamma^2 + \gamma - 1 = 0 \quad (9.22)$$

(见练习 9.7)。

为满足定理 9.1 的条件(i),式(9.22)的解 γ 必须为正。注意到 $f(0) = -1$ 和 $f(+\infty) = +\infty$,则方程至少存在一个正的根 γ 。此外,由笛卡儿正负号法则可知,式(9.22)最多存在 3 个实根。

下面给出选择不同 s 时的 γ 值:

| | | | | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| s | 0.09 | 0.08 | 0.07 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 | 0.01 |
| γ | 0.755 | 0.778 | 0.804 | 0.835 | 0.873 | 0.919 | 0.979 | 1.065 | 1.211 |

练习

9.1 考虑二维实时跟踪系统:

$$\begin{cases} \mathbf{x}_{k+1} = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \xi_k \\ v_k = [1 \ 0] \mathbf{x}_k + \eta_k \end{cases}$$

式中, $h > 0$, $\{\xi_k\}$ 、 $\{\eta_k\}$ 是不相关的零均值高斯白噪声序列。与该系统相关的 α - β 跟踪器定义为

$$\begin{cases} \dot{\mathbf{x}}_k = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \dot{\mathbf{x}}_{k-1} + \begin{bmatrix} \alpha \\ \beta/h \end{bmatrix} (v_k - [1 \ 0] \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \dot{\mathbf{x}}_{k-1}) \\ \dot{\mathbf{x}}_0 = E(\mathbf{x}_0) \end{cases}$$

- (a) 推导该 α - β 跟踪器的解耦卡尔曼滤波算法;
- (b) 给出该 α - β 跟踪器是极限卡尔曼滤波器的条件。

9.2 验证 9.2 节的实时跟踪系统(9.11)的 x_k 、 \dot{x}_k 和 \ddot{x}_k 的解耦公式。

9.3 考虑三维雷达跟踪系统:

$$\begin{cases} \dot{\mathbf{x}}_{k+1} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \underline{\xi}_k \\ v_k = [1 \ 0 \ 0] \mathbf{x}_k + w_k \end{cases}$$

其中 $\{w_k\}$ 是有色噪声序列, 定义为 $w_k = s w_{k-1} + \eta_k$, $\{\underline{\xi}_k\}, \{\eta_k\}$ 与第5章定义的一致, 为不相关零均值高斯白噪声序列。该系统相应的 α - β - γ 跟踪器通过下面的算法定义:

$$\begin{cases} \dot{\mathbf{X}}_k = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & s \end{bmatrix} \check{\mathbf{X}}_{k-1} + \begin{bmatrix} \alpha \\ \beta/h \\ \gamma/h^2 \\ \theta \end{bmatrix} \left\{ v_k - [1 \ 0 \ 0] \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & s \end{bmatrix} \check{\mathbf{X}}_{k-1} \right\} \\ \check{\mathbf{X}}_0 = \begin{bmatrix} E(\mathbf{x}_0) \\ \mathbf{0} \end{bmatrix} \end{cases}$$

式中 α, β, γ 和 θ 是常值。(见图9.2)

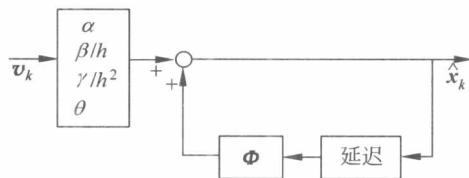


图9.2 三维雷达跟踪系统

(a) 计算矩阵

$$\Phi = \left\{ \mathbf{I} - \begin{bmatrix} \alpha \\ \beta/h \\ \gamma/h^2 \\ \theta \end{bmatrix} [1 \ 0 \ 0] \right\} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & s \end{bmatrix}$$

(b) 应用克莱姆法则, 求解系统

$$[\mathbf{zI} - \Phi] \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \\ \tilde{X}_3 \\ W \end{bmatrix} = \mathbf{z} \begin{bmatrix} \alpha \\ \beta/h \\ \gamma/h^2 \\ \theta \end{bmatrix} V$$

中的 $\tilde{X}_1, \tilde{X}_2, \tilde{X}_3$ 和 W (由 $\alpha-\beta-\gamma-\theta$ 濾波器的 z -变换, 可以得到上面的系统);

(c) 由 $\tilde{X}_1, \tilde{X}_2, \tilde{X}_3$ 和 W 的 z -反变换, 给出 $\alpha-\beta-\gamma-\theta$ 濾波器的解耦滤波方程;

(d) 验证当有色噪声 $\{\eta_k\}$ 变为白噪声, 即 $s=0, \theta=0$ 时, (c) 的解耦滤波方程可简化为在 9.2 节得到的形式, 其中 $g_1=\alpha, g_2=\beta/h$ 和 $g_3=\gamma/h^2$ 。

9.4 验证方程(9.17)~(9.20)。

9.5 证明定理 9.1, 且条件(i)~(iii)与采样时间 h 无关。

9.6 验证方程(9.21)。

9.7 验证方程(9.22)。

第 10 章

区间系统的卡尔曼滤波

如果系统的一些参数,如系统矩阵的某些元素,不能确定或随时间改变,这时卡尔曼滤波算法就不能直接应用。在这种情况下,就需要具有处理不确定性的鲁棒卡尔曼滤波。本章将介绍一种鲁棒卡尔曼滤波算法。

考虑系统:

$$\begin{cases} \underline{\mathbf{x}}_{k+1} = \mathbf{A}_k \underline{\mathbf{x}}_k + \boldsymbol{\Gamma}_k \underline{\boldsymbol{\xi}}_k \\ \underline{\mathbf{v}}_k = \mathbf{C}_k \underline{\mathbf{x}}_k + \underline{\boldsymbol{\eta}}_k \end{cases} \quad (10.1)$$

式中 \mathbf{A}_k 、 $\boldsymbol{\Gamma}_k$ 和 \mathbf{C}_k 分别是已知的 $n \times n$ 、 $n \times p$ 和 $q \times n$ 阶矩阵, $1 \leq p, q \leq n$ 。且

$$E(\underline{\boldsymbol{\xi}}_k) = \mathbf{0}, \quad E(\underline{\boldsymbol{\xi}}_k \underline{\boldsymbol{\xi}}_l^T) = \mathbf{Q}_k \delta_{kl}$$

$$E(\underline{\boldsymbol{\eta}}_k) = \mathbf{0}, \quad E(\underline{\boldsymbol{\eta}}_k \underline{\boldsymbol{\eta}}_l^T) = \mathbf{R}_k \delta_{kl}$$

$$E(\underline{\boldsymbol{\xi}}_k \boldsymbol{\eta}_l^T) = \mathbf{0}, \quad E(\underline{\mathbf{x}}_0 \underline{\boldsymbol{\xi}}_k^T) = \mathbf{0}, \quad E(\underline{\mathbf{x}}_0 \underline{\boldsymbol{\eta}}_k^T) = \mathbf{0}$$

对所有的 $k, l = 0, 1, \dots, Q_k$ 和 R_k 是正定对称矩阵。

如果所有的常值矩阵 \mathbf{A}_k 、 $\boldsymbol{\Gamma}_k$ 和 \mathbf{C}_k 都已知, 卡尔曼滤波就能应用到式(10.1)所示的系统。这样递推地使用量测数据 $\{\underline{\mathbf{v}}_k\}$, 就可以得到未知状态向量 $\{\underline{\mathbf{x}}_k\}$ 的最优估计值 $\{\hat{\underline{\mathbf{x}}}_k\}$ 。然而, 如果系统矩阵中的某些元素是未知的或不确定的, 就必须对整个滤波过程进行修改。假设仅知道所有未知参数的范围, 则对于 $k=0, 1, \dots$, 可记

$$\mathbf{A}_k^I = \mathbf{A}_k + \Delta \mathbf{A}_k = [\mathbf{A}_k - |\Delta \mathbf{A}_k|, \mathbf{A}_k + |\Delta \mathbf{A}_k|]$$

$$\boldsymbol{\Gamma}_k^I = \boldsymbol{\Gamma}_k + \Delta \boldsymbol{\Gamma}_k = [\boldsymbol{\Gamma}_k - |\Delta \boldsymbol{\Gamma}_k|, \boldsymbol{\Gamma}_k + |\Delta \boldsymbol{\Gamma}_k|]$$

$$\mathbf{C}_k^l = \mathbf{C}_k + \Delta \mathbf{C}_k = [\mathbf{C}_k - |\Delta \mathbf{C}_k|, \mathbf{C}_k + |\Delta \mathbf{C}_k|]$$

式中, $|\Delta \mathbf{A}_k|$ 、 $|\Delta \boldsymbol{\Gamma}_k|$ 和 $|\Delta \mathbf{C}_k|$ 是固定的不确定范围。

对于 $k=0,1,\dots$, 相应的系统

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k^l \mathbf{x}_k + \boldsymbol{\Gamma}_k^l \underline{\boldsymbol{\xi}}_k \\ \mathbf{v}_k = \mathbf{C}_k^l \mathbf{x}_k + \underline{\boldsymbol{\eta}}_k \end{cases} \quad (10.2)$$

称为区间系统。

在此框架下, 如何改进原始的卡尔曼滤波算法并应用到式(10.2)的区间系统上, 这一问题将在本章内讨论。

10.1 区间数学

本节将给出在本章要用到的区间算术和区间分析的基本结论。

10.1.1 区间及其特性

定义在 $R=(-\infty, \infty)$ 上的有界闭子集 $[\underline{x}, \bar{x}]$ 称为是一个区间。在特殊情况下, 一个单独的点 $x \in R$ 也可以被认为是一个退化了的区间, 其 $\underline{x}=\bar{x}=x$ 。

下面说明一些关于区间的概念和特性。

(a) 相等: 当且仅当 $\underline{x}_1=\underline{x}_2, \bar{x}_1=\bar{x}_2$ 时, 两个区间 $[\underline{x}_1, \bar{x}_1]$ 和 $[\underline{x}_2, \bar{x}_2]$ 相等, 可以表示为

$$[\underline{x}_1, \bar{x}_1] = [\underline{x}_2, \bar{x}_2]$$

(b) 交: 两个区间 $[\underline{x}_1, \bar{x}_1]$ 和 $[\underline{x}_2, \bar{x}_2]$ 的交可以表示为

$$[\underline{x}_1, \bar{x}_1] \cap [\underline{x}_2, \bar{x}_2] = [\max\{\underline{x}_1, \underline{x}_2\}, \min\{\bar{x}_1, \bar{x}_2\}]$$

当且仅当 $\underline{x}_1 > \bar{x}_2$ 或 $\underline{x}_2 > \bar{x}_1$ 时, 这两个区间不相交, 表示为

$$[\underline{x}_1, \bar{x}_1] \cap [\underline{x}_2, \bar{x}_2] = \emptyset$$

(c) 并: 两个相交区间 $[\underline{x}_1, \bar{x}_1]$ 和 $[\underline{x}_2, \bar{x}_2]$ 的并可以表示为

$$[\underline{x}_1, \bar{x}_1] \cup [\underline{x}_2, \bar{x}_2] = [\min\{\underline{x}_1, \underline{x}_2\}, \max\{\bar{x}_1, \bar{x}_2\}]$$

注意只有两个区间相交时, 并才有意义, 也就是

$$[\underline{x}_1, \bar{x}_1] \cap [\underline{x}_2, \bar{x}_2] \neq \emptyset$$

否则是没有意义的, 因为其结果不是一个区间。

(d) 不等式: 当且仅当 $\underline{x}_1 < \underline{x}_2$ (或 $\bar{x}_1 > \bar{x}_2$), 区间 $[\underline{x}_1, \bar{x}_1]$ 小于(或大于)区间 $[\underline{x}_2, \bar{x}_2]$, 并且表示为

$$[\underline{x}_1, \bar{x}_1] < [\underline{x}_2, \bar{x}_2] \text{ (或 } [\underline{x}_1, \bar{x}_1] > [\underline{x}_2, \bar{x}_2])$$

否则就不能对它们进行比较。注意在区间中 \leq 以及 \geq 是没有定义的。

(e) 包含: 当且仅当 $\underline{x}_2 \leq \underline{x}_1$ 且 $\bar{x}_1 \leq \bar{x}_2$, 即当且仅当 $[\underline{x}_1, \bar{x}_1]$ 是 $[\underline{x}_2, \bar{x}_2]$ 的子集(子空间)时, 区间 $[\underline{x}_1, \bar{x}_1]$ 被包含在区间 $[\underline{x}_2, \bar{x}_2]$ 中, 表示为

$$[\underline{x}_1, \bar{x}_1] \subseteq [\underline{x}_2, \bar{x}_2]$$

例如, 有三个给定的区间 $X_1 = [-1, 0]$, $X_2 = [-1, 2]$ 和 $X_3 = [2, 10]$, 则

$$X_1 \cap X_2 = [-1, 0] \cap [-1, 2] = [-1, 0]$$

$$X_1 \cap X_3 = [-1, 0] \cap [2, 10] = \emptyset$$

$$X_2 \cap X_3 = [-1, 2] \cap [2, 10] = [2, 2] = 2$$

$$X_1 \cup X_2 = [-1, 0] \cup [-1, 2] = [-1, 2]$$

$$X_1 \cup X_3 = [-1, 0] \cup [2, 10] \text{ 没有定义}$$

$$X_2 \cup X_3 = [-1, 2] \cup [2, 10] = [-1, 10]$$

$$X_1 = [-1, 0] < [2, 10] = X_3$$

$$X_1 = [-1, 0] \subset [-1, 10] = X_2$$

10.1.2 区间运算

设 $[\underline{x}, \bar{x}]$, $[\underline{x}_1, \bar{x}_1]$ 和 $[\underline{x}_2, \bar{x}_2]$ 是区间, 区间的基本运算操作可以定义如下。

(a) 加

$$[\underline{x}_1, \bar{x}_1] + [\underline{x}_2, \bar{x}_2] = [\underline{x}_1 + \underline{x}_2, \bar{x}_1 + \bar{x}_2]$$

(b) 减

$$[\underline{x}_1, \bar{x}_1] - [\underline{x}_2, \bar{x}_2] = [\underline{x}_1 - \bar{x}_2, \bar{x}_1 - \underline{x}_2]$$

(c) 逆

$$\begin{cases} [\underline{x}, \bar{x}]^{-1} = [1/\bar{x}, 1/\underline{x}], & 0 \notin [\underline{x}, \bar{x}] \\ [\underline{x}, \bar{x}]^{-1} \text{ 没有定义,} & 0 \in [\underline{x}, \bar{x}] \end{cases}$$

(d) 乘

$$[\underline{x}_1, \bar{x}_1] \cdot [\underline{x}_2, \bar{x}_2] = [\underline{y}, \bar{y}]$$

式中

$$\underline{y} = \min\{\underline{x}_1 \underline{x}_2, \underline{x}_1 \bar{x}_2, \bar{x}_1 \underline{x}_2, \bar{x}_1 \bar{x}_2\}$$

$$\bar{y} = \max\{\underline{x}_1 \underline{x}_2, \underline{x}_1 \bar{x}_2, \bar{x}_1 \underline{x}_2, \bar{x}_1 \bar{x}_2\}$$

(e) 除

$$\begin{cases} [\underline{x}_1, \bar{x}_1] / [\underline{x}_2, \bar{x}_2] = [\underline{x}_1, \bar{x}_1] \cdot [\underline{x}_2, \bar{x}_2]^{-1}, & 0 \notin [\underline{x}_2, \bar{x}_2] \\ \text{无意义,} & 0 \in [\underline{x}_2, \bar{x}_2] \end{cases}$$

对于三个区间 $X = [\underline{x}, \bar{x}]$ 、 $Y = [\underline{y}, \bar{y}]$ 和 $Z = [\underline{z}, \bar{z}]$ ，考虑区间的加(+)、减(-)、乘(\cdot)和除(/)操作，也就是

$$Z = X * Y, \quad * \in \{+, -, \cdot, /\}$$

很明显 $X * Y$ 也是一个区间。换句话说，区间在四种代数运算操作 $\{+, -, \cdot, /\}$ 下是封闭的。同理，实数 $x, y, z \dots$ 和退化区间 $[x, x], [y, y], [z, z] \dots$ 是相同的。所以可以简单地将点区间操作 $[x, x] * Y$ 表示为 $x * Y$ 。此外，为了表示的简便，乘法运算符号“ \cdot ”通常可以省略。

与常规算术类似，区间算术遵从以下的基本算术性质（见练习 10.1）：

$$X + Y = Y + X$$

$$Z + (X + Y) = (Z + X) + Y$$

$$XY = YX$$

$$Z(XY) = (ZX)Y$$

$$X + 0 = 0 + X = X \quad X \cdot 0 = 0 \cdot X = 0, \text{ 其中 } 0 = [0, 0]$$

$$XI = IX = X, \text{ 其中 } I = [1, 1]$$

$Z(X+Y) \subseteq ZX+ZY$, 其中“=”只有下列条件之一满足时才成立：

$$(a) Z = [z, z]$$

$$(b) X = Y = 0$$

$$(c) \text{对所有的 } x \in X \text{ 和 } y \in Y, xy \geq 0$$

此外，下面的区间操作重要特性被称为单调包含特性。

定理 10.1 设区间 X_1, X_2, Y_1 和 Y_2 ，有

$$X_1 \subseteq Y_1, \quad X_2 \subseteq Y_2$$

则对于任意操作 $* \in \{+, -, \cdot, /\}$ ，都有

$$X_1 * X_2 \subseteq Y_1 * Y_2$$

此特性可以由 $X_1 \subseteq Y_1$ 和 $X_2 \subseteq Y_2$ 直接得出，即

$$\begin{aligned} X_1 * X_2 &= \{x_1 * x_2 \mid x_1 \in X_1, x_2 \in X_2\} \\ &\subseteq \{y_1 * y_2 \mid y_1 \in Y_1, y_2 \in Y_2\} \\ &= Y_1 * Y_2 \end{aligned}$$

推论 10.1 设有区间 X 和 Y ，且有 $x \in X$ 和 $y \in Y$ ，则

$x * y \subseteq X * Y$, 对所有的 $* \in \{+, -, \cdot, /\}$

从表面上看, 上面的定理和推论对有些运算, 如求逆、减和除, 好像不满足这样的单调包含特性。然而, 除了上面的证明, 考虑关于区间 $X = [0.2, 0.4]$ 和 $Y = [0.1, 0.5]$ 的简单例子。很明显, $X \subseteq Y$ 。首先来看 $\frac{I}{X} \subseteq \frac{I}{Y}$, 其中 $I = [1.0, 1.0]$, 事实上,

$$\frac{1}{X} = \frac{[1.0, 1.0]}{[0.2, 0.4]} = [2.5, 5.0] \quad \frac{1}{Y} = \frac{[1.0, 1.0]}{[0.1, 0.5]} = [2.0, 10.0]$$

通过计算得

$$I - X = [1.0, 1.0] - [0.2, 0.4] = [0.6, 0.8]$$

和

$$I - Y = [1.0, 1.0] - [0.1, 0.5] = [0.5, 0.9]$$

还可以验证 $I - X \subseteq I - Y$ 。

更进一步, 将上面两个操作组合:

$$\frac{I}{I - X} = [5/4, 5/3], \quad \frac{I}{I - Y} = [10/9, 2]$$

依然可得 $\frac{I}{I - X} \subseteq \frac{I}{I - Y}$

下面将区间的概念和区间算术扩展到区间向量和区间矩阵。区间向量和区间矩阵有着相似的定义。例如:

$$\mathbf{A}^I = \begin{bmatrix} [2, 3] & [0, 1] \\ [1, 2] & [2, 3] \end{bmatrix}, \quad \mathbf{b}^I = \begin{bmatrix} [0, 10] \\ [-6, 1] \end{bmatrix}$$

分别是区间矩阵和区间向量。

设 $\mathbf{A}^I = [a_{ij}^I]$, $\mathbf{B}^I = [b_{ij}^I]$ 是 $n \times m$ 阶区间矩阵。如果对于所有的 $i = 1, \dots, n$ 和 $j = 1, 2, \dots, m$, 都有 $a_{ij}^I = b_{ij}^I$, 则说 \mathbf{A}^I 和 \mathbf{B}^I 相等。如果对于所有的 $i = 1, \dots, n$ 和 $j = 1, 2, \dots, m$, 都有 $a_{ij}^I \subseteq b_{ij}^I$, 则说 \mathbf{A}^I 包含于 \mathbf{B}^I , 表示为 $\mathbf{A}^I \subseteq \mathbf{B}^I$, 尤其当 $\mathbf{A}^I = \mathbf{A}$ 是普通的常值矩阵时, 表示为 $\mathbf{A} \in \mathbf{B}^I$ 。

区间矩阵的基本操作包括如下方面。

(a) 加法和减法

$$\mathbf{A}^I \pm \mathbf{B}^I = [a_{ij}^I \pm b_{ij}^I]$$

(b) 乘法: 对于两个 $n \times r$ 和 $r \times m$ 矩阵 \mathbf{A}^I 和 \mathbf{B}^I ,

$$\mathbf{A}^I \mathbf{B}^I = \left[\sum_{k=1}^r a_{ik}^I b_{kj}^I \right]$$

(c) 求逆: 对于 $n \times n$ 阶区间矩阵 \mathbf{A}^I , 其行列式的值 $\det[\mathbf{A}^I] \neq 0$, 则

$$[\mathbf{A}^I]^{-1} = \frac{\text{adj}[\mathbf{A}^I]}{\det[\mathbf{A}^I]}$$

比如, $\mathbf{A}^I = \begin{bmatrix} [2,3] & [0,1] \\ [1,2] & [2,3] \end{bmatrix}$, 则

$$\begin{aligned} [\mathbf{A}^I]^{-1} &= \frac{\text{adj}[\mathbf{A}^I]}{\det[\mathbf{A}^I]} \\ &= \frac{\begin{bmatrix} [2,3] & -[0,1] \\ -[1,2] & [2,3] \end{bmatrix}}{[2,3][2,3] - [0,1][1,2]} \\ &= \begin{bmatrix} [2/9, 3/2] & [-1/2, 0] \\ [-1, -1/9] & [2/9, 3/2] \end{bmatrix} \end{aligned}$$

区间矩阵(包括区间向量)和区间一样, 服从许多算术运算规则(见练习 10.2)。

10.1.3 有理区间函数

设 S_1 和 S_2 是 \mathbb{R} 上的区间, $f: S_1 \rightarrow S_2$ 是普通的单变量(如, 点到点)实值函数, 分别用 \sum_{S_1} 和 \sum_{S_2} 表示 S_1 和 S_2 的所有子区间族。区间到区间的函数 $f^l: \sum_{S_1} \rightarrow \sum_{S_2}$, 定义

$$f^l(X) = \{f(x) \in S_2 : x \in X, X \in \sum_{S_1}\}$$

称为点到点函数 f 在区间 S_1 上一致扩展。很明显, 其值域是

$$f^l(X) = \bigcup_{x \in X} \{f(x)\}$$

这是包含单点 $f(x)$ ($x \in X$) 的 S_2 所有子集的集合。

一致扩展 $f^l: \sum_{S_1} \rightarrow \sum_{S_2}$ 的特性可以直接从定义得到, 即

$$X, Y \in \sum_{S_1} \text{ 和 } X \subseteq Y \Rightarrow f^l(X) \subseteq f^l(Y)$$

通常, n 维变量 X_1, \dots, X_n 的区间到区间的函数 F 被称为具有单调包含特性, 如果

$$X_i \subseteq Y_i, \quad \forall i = 1, \dots, n \Rightarrow F(X_1, \dots, X_n) \subseteq F(Y_1, \dots, Y_n)$$

注意, 并不是所有的区间到区间的函数都具有这样的特性。

然而, 所有的致扩展都具有单调包含特性, 因为对于加、减、乘和除 (+, -, *, /), 区间算术函数是实值算术函数的一致扩展。正如前面讨论的(见定理 10.1 和推论 10.1), 区间算术具有单调包含特性。

一个区间到区间的函数就简称为区间函数。区间向量和区间矩阵具有相似的定义。一个区间函数称为有理区间函数, 如果它的值是由有限个区间算术运

算定义的。有理区间函数的例子如：对于区间 X, Y 和 Z , $X+Y^2+Z^3$ 和 $(X^2+Y^2)/Z$ 等，对于后一个，假设 $0 \notin Z$ 。

根据“ \subseteq ”的传递性，可以得到所有的有理区间函数都具有单调包含特性。这可以通过数学归纳法证明。

下面，设 $f=f(x_1, \dots, x_n)$ 是一普通 n 维变量实值函数， X_1, \dots, X_n 是区间。如果

$$F(x_1, \dots, x_n) = f(x_1, \dots, x_n), \quad \forall x_i \in X_i, \quad i = 1, \dots, n$$

则区间函数 $F=F(X_1, \dots, X_n)$ 称为是 f 的区间扩展。同时注意并非所有的区间扩展都具有单调包含特性。

可以证明下面的结论(见练习 10.3)。

定理 10.2 如果 F 是 f 的区间扩展，并且具有单调包含特性，则 f 的一致扩展 f^I 满足：

$$f^I(X_1, \dots, X_n) \subseteq F(X_1, \dots, X_n)$$

因为有理区间函数具有单调包含特性，故有：

推论 10.2 如果 F 是有理区间函数，并且是 f 的区间扩展，则

$$f^I(X_1, \dots, X_n) \subseteq F(X_1, \dots, X_n)$$

这个推论针对定义在 \mathbf{R}^n 上的一个 n 维普通有理函数，给出了函数值的上界和下界的极限值。

现在举一个具有单调包含特性的实值区间函数的例子。考虑计算下面的函数：

$$f^I(X, A) = \frac{AX}{I-X}$$

对于 $X_1=[2,3]$ 、 $A_1=[0,2]$ 和 $X_2=[2,4]$ 、 $A_2=[0,3]$ 两种情况，这里 $X_1 \subset X_2$ 、 $A_1 \subset A_2$ ，通过直接计算可以得到

$$f_1^I(X_1, A_1) = \frac{[0,2] \cdot [2,3]}{[1,1] - [2,3]} = [-6,0]$$

$$f_2^I(X_2, A_2) = \frac{[0,3] \cdot [2,4]}{[1,1] - [2,4]} = [-12,0]$$

与期望的一致，确实得到了 $f_1^I(X_1, A_1) \subset f_2^I(X_2, A_2)$ 。

最后，基于推论 10.2，当 X^I 不包含零时，对于形如 X^I/X^I 的区间除法，可以首先检验与其对应的普通函数的运算可以得到 $x/x=1$ ，然后回到区间集合得到最后的结论。所以，作为区间计算中的一个习惯，对于不包含零的区间 X^I ，可以象征性地写为 $X^I/X^I=1$ 。

10.1.4 区间期望和方差

设 $f(x)$ 是定义在某个区间 X 上的普通函数。对于一些不依赖于 $x, y \in X$ 的正的常数 L ，如果 f 满足通常的利普希茨(Lipschitz)条件：

$$|f(x) - f(y)| \leq L |x - y|$$

则函数 f 的一致扩展 f^l 称为 f 在区间 X 上的利普希茨区间扩展。

设 $B(X)$ 是定义在 X 上的一类在数值计算中经常使用的函数, 例如四种算术函数($+$, $-$, \cdot , $/$)以及初等函数如 $e^{(\cdot)}$, $\ln(\cdot)$, $\sqrt{\cdot}$ 等。本章中将仅使用这些常用的基本函数中的一部分, $B(X)$ 只是为了表示方便。

设 N 是一个正整数。将区间 $[a, b] \subseteq X$ 细分为 N 个子区间, $X_1 = [\underline{X}_1, \bar{X}_1], \dots, X_N = [\underline{X}_N, \bar{X}_N]$, 使得

$$a = \underline{X}_1 < \bar{X}_1 = \underline{X}_2 < \bar{X}_2 = \dots = \underline{X}_N < \bar{X}_N = b$$

此外, 对于任意 $f \in B(X)$, 设 F 是 f 的一个定义在所有 $X_i (i=1, \dots, N)$ 上的利普希茨区间扩展, 并假设 F 满足单调包含特性, 记

$$S_N(F; [a, b]) = \frac{b-a}{N} \sum_{i=1}^N F(X_i)$$

则有

$$\int_a^b f(t) dt = \bigcap_{N=1}^{\infty} S_N(F; [a, b]) = \lim_{N \rightarrow \infty} S_N(F; [a, b])$$

注意到, 如果递推地定义

$$\begin{cases} Y_1 = S_1 \\ Y_{k+1} = S_{k+1} \cap Y_k, \quad k = 1, 2, \dots \end{cases}$$

式中, $S_k = S_k(F; [a, b])$, 则 $\{Y_k\}$ 是收敛到积分 $\int_a^b f(t) dt$ 真实值的区间套序列。

注意到这里使用的利普希茨区间扩展 F 具有以下特性, 即对于任何实数 $x \in \mathbf{R}$, $F(x)$ 都是一个实数。然而, 对于其他具有单调包含特性但不是利普希茨的区间函数, 即使 x 是实数, 其相应的函数 $F(x)$ 也可能有区间函数。

基于前面介绍的区间数学, 接下来介绍几个重要的概念。

设 X 是具有实值随机变量的区间, 并设

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}, \quad x \in X$$

是具有已知 $\mu_x, \sigma_x > 0$ 的普通高斯密度函数。则 $f(x)$ 具有利普希茨区间扩展。定义区间期望为

$$E(X) = \int_{-\infty}^{\infty} xf(x) dx = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma_x} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} dx, \quad x \in X \quad (10.3)$$

以及区间方差为

$$\begin{aligned}\text{Var}(X) &= E([X - E(X)]^2) = \int_{-\infty}^{+\infty} (x - \mu_x)^2 f(x) dx \\ &= \int_{-\infty}^{\infty} \frac{(x - \mu_x)^2}{\sqrt{2\pi}\sigma_x} \exp\left\{-\frac{(x - \mu_x)^2}{2\sigma_x^2}\right\} dx \quad x \in X,\end{aligned}\quad (10.4)$$

容易验证,当 $a \rightarrow -\infty$ 和 $b \rightarrow \infty$ 时,从前面定义的定积分便可以得到这两个定义。

同理,对于另一实值随机变量区间 Y 定义条件区间期望

$$\begin{aligned}E(X | y \in Y) &= \int_{-\infty}^{\infty} xf(x | y) dx = \int_{-\infty}^{\infty} x \frac{f(x, y)}{f(y)} dx \\ &= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma_{xy}} \exp\left\{-\frac{(x - \mu_{xy})^2}{2\sigma_{xy}^2}\right\} dx, \quad x \in X\end{aligned}\quad (10.5)$$

和条件方差

$$\begin{aligned}\text{Var}(X | y \in Y) &= E((x - \mu_x)^2 | y \in Y) \\ &= \int_{-\infty}^{\infty} [x - E(x | y \in Y)]^2 f(x | y) dx \\ &= \int_{-\infty}^{\infty} [x - E(x | y \in Y)]^2 \frac{f(x, y)}{f(y)} dx \\ &= \int_{-\infty}^{\infty} \frac{[x - E(x | y \in Y)]^2}{\sqrt{2\pi}\tilde{\sigma}} \exp\left\{-\frac{(x - \tilde{\mu})^2}{2\tilde{\sigma}^2}\right\} dx, \quad x \in X\end{aligned}\quad (10.6)$$

同理,可以通过定义的区间除法操作(注意零不包含在高斯区间密度函数的分母中)验证上述两个定义。

在上面的定义中,

$$\tilde{\mu} = \mu_x + \sigma_{xy}^2 (y - \mu_y) / \sigma_y^2, \quad \tilde{\sigma}^2 = \sigma_x^2 - \sigma_{xy}^2 \sigma_{yx}^2 / \sigma_y^2$$

式中

$$\sigma_{xy}^2 = \sigma_{yx}^2 = E(XY) - E(X)E(Y) = E(xy) - E(x)E(y), \quad x \in X$$

还可以验证(见练习 10.4):

$$E(X | y \in Y) = E(x) + \sigma_{xy}^2 [y - E(y)] / \sigma_y^2, \quad x \in X \quad (10.7)$$

$$\text{Var}(X | y \in Y) = \text{Var}(x) - \sigma_{xy}^2 \sigma_{yx}^2 / \sigma_y^2, \quad x \in X \quad (10.8)$$

所有这些量都是典型的有理区间函数,所以可以应用推论 10.2。

10.2 区间卡尔曼滤波

现在回到(10.2)所示的区间系统。可以看出该系统有一个由区间矩阵的所有上界元素定义的上界系统:

$$\begin{cases} \mathbf{x}_{k+1} = [\mathbf{A}_k + |\Delta\mathbf{A}_k|] \mathbf{x}_k + [\boldsymbol{\Gamma}_k + |\Delta\boldsymbol{\Gamma}_k|] \underline{\boldsymbol{\xi}}_k \\ \mathbf{v}_k = [\mathbf{C}_k + |\Delta\mathbf{C}_k|] \mathbf{x}_k + \underline{\boldsymbol{\eta}}_k \end{cases} \quad (10.9)$$

及一个由区间矩阵的所有下界元素确定的下界系统：

$$\begin{cases} \mathbf{x}_{k+1} = [\mathbf{A}_k - |\Delta\mathbf{A}_k|] \mathbf{x}_k + [\boldsymbol{\Gamma}_k - |\Delta\boldsymbol{\Gamma}_k|] \underline{\boldsymbol{\xi}}_k \\ \mathbf{v}_k = [\mathbf{C}_k - |\Delta\mathbf{C}_k|] \mathbf{x}_k + \underline{\boldsymbol{\eta}}_k \end{cases} \quad (10.10)$$

但对这两个边界系统应用标准卡尔曼滤波算法，得到的两个滤波轨迹不一定能包围区间系统(10.2)的所有可能最优解(见练习 10.5)。事实上，在这两个边界轨迹和所有最优滤波解的集合间并没有确定的关系：因为噪声的干扰，这两个边界轨迹和它们相邻的轨迹常常会互相交叉。所以迫切需要一种能够提供包含区间系统所有估计值轨迹的新滤波算法。下面将要介绍的区间卡尔曼滤波方案能够满足这样的需求。

10.2.1 区间卡尔曼滤波方案

回顾在第3章介绍的标准卡尔曼滤波算法的推导过程，其中只应用到了矩阵代数运算(加、减、乘、除)和(条件)期望和方差。由前面讨论可知，所有这些运算在区间矩阵和有理区间函数中都有定义，所以可以采用相同的方式来推导区间系统(10.2)的卡尔曼滤波算法。该区间卡尔曼滤波算法简单介绍如下：

区间卡尔曼滤波算法

主过程

$$\begin{aligned} \hat{\mathbf{x}}_0^I &= \mathbf{E}(\mathbf{x}_0^I) \\ \hat{\mathbf{x}}_k^I &= \mathbf{A}_{k-1}^I \hat{\mathbf{x}}_{k-1}^I + \mathbf{G}_k^I [\mathbf{v}_k^I - \mathbf{C}_k^I \mathbf{A}_{k-1}^I \hat{\mathbf{x}}_{k-1}^I] \\ k &= 1, 2, \dots \end{aligned} \quad (10.11)$$

从过程

$$\begin{aligned} \mathbf{P}_0^I &= \text{Var}(\mathbf{x}_0^I) \\ \mathbf{M}_{k-1}^I &= \mathbf{A}_{k-1}^I \mathbf{P}_{k-1}^I [\mathbf{A}_{k-1}^I]^T + \mathbf{B}_{k-1}^I \mathbf{Q}_{k-1} [\mathbf{B}_{k-1}^I]^T \\ \mathbf{G}_k^I &= \mathbf{M}_{k-1}^I [\mathbf{C}_k^I]^T [[\mathbf{C}_k^I] \mathbf{M}_{k-1}^I [\mathbf{C}_k^I]^T + \mathbf{R}_k]^{-1} \\ \mathbf{P}_k^I &= [\mathbf{I} - \mathbf{G}_k^I \mathbf{C}_k^I] \mathbf{M}_{k-1}^I [\mathbf{I} - \mathbf{G}_k^I \mathbf{C}_k^I]^T + [\mathbf{G}_k^I] \mathbf{R}_k [\mathbf{G}_k^I]^T \\ k &= 1, 2, \dots \end{aligned} \quad (10.12)$$

将该算法和式(3.25)的标准卡尔曼滤波算法比较，可以发现除了所有的矩阵和向量在算法(10.11)~(10.12)中为区间，它们在形式上是严格相似的。要注意，区间

估计轨迹可能会快速地分开。但这归咎于保守的区间建模,而不是新的滤波算法。

需要注意的是,通过理论分析该区间卡尔曼滤波算法对于区间系统(10.2)是最优的,和标准卡尔曼滤波算法一样,因为在这个推导过程中没有任何近似。区间卡尔曼滤波算法的滤波结果是区间估计的序列 $\{\hat{x}_k^l\}$,它包含了区间系统可能产生的状态 $\{x_k\}$ 的所有可能的最优估计 $\{\hat{x}_k\}$ 。因此,为了使得区间估计的结果能包含所有可能的最优解,其范围通常比较宽。该区间卡尔曼滤波算法的滤波结果包含了一切可能的结果,但是比较保守。

同样应该注意到类似于通常情况下的随机向量(量测数据) v_k ,在前面的区间卡尔曼滤波算法中,区间数据向量 v_k^l 在实现之前是个不确定的区间向量(例如,在该数据实际获取之前),但是在其被测量并获得之后,就成为一个普通的常值向量。这在算法的实现过程中应该避免混淆。

10.2.2 次优区间卡尔曼滤波

为了改进区间卡尔曼滤波算法(10.11)~(10.12)的计算效率,需要做一些适当的近似。在本节中,将介绍一种次优区间卡尔曼滤波方案,把区间矩阵的逆用其最坏逆代替,而保持其他所有的项不变。

设

$$\mathbf{C}_k^l = \mathbf{C}_k + \Delta\mathbf{C}_k, \quad \mathbf{M}_{k-1}^l = \mathbf{M}_{k-1} + \Delta\mathbf{M}_{k-1}$$

其中 \mathbf{C}_k 是 \mathbf{C}_k^l 的中点, \mathbf{M}_{k-1} 是 \mathbf{M}_{k-1}^l 的中点(例如,区间矩阵的标称值)。

记

$$\begin{aligned} [[\mathbf{C}_k^l] \mathbf{M}_{k-1}^l [\mathbf{C}_k^l]^T + \mathbf{R}_k]^{-1} &= [[\mathbf{C}_k + \Delta\mathbf{C}_k][\mathbf{M}_{k-1} + \Delta\mathbf{M}_{k-1}] [\mathbf{C}_k + \Delta\mathbf{C}_k]^T + \mathbf{R}_k]^{-1} \\ &= [\mathbf{C}_k \mathbf{M}_{k-1} \mathbf{C}_k^T + \Delta\mathbf{R}_k]^{-1} \end{aligned}$$

式中

$$\begin{aligned} \Delta\mathbf{R}_k &= \mathbf{C}_k \mathbf{M}_{k-1} [\Delta\mathbf{C}_k]^T + \mathbf{C}_k [\Delta\mathbf{M}_{k-1}] \mathbf{C}_k^T + \mathbf{C}_k [\Delta\mathbf{M}_{k-1}] [\Delta\mathbf{C}_k]^T \\ &\quad + [\Delta\mathbf{C}_k] \mathbf{M}_{k-1} \mathbf{C}_k^T + [\Delta\mathbf{C}_k] \mathbf{M}_{k-1} [\Delta\mathbf{C}_k]^T + [\Delta\mathbf{C}_k] [\Delta\mathbf{M}_{k-1}] \mathbf{C}_k^T \\ &\quad + [\Delta\mathbf{C}_k] [\Delta\mathbf{M}_{k-1}] [\Delta\mathbf{C}_k]^T + \mathbf{R}_k \end{aligned}$$

这样,在算法(10.11)~(10.12)中,将 $\Delta\mathbf{R}_k$ 用其上界矩阵 $|\Delta\mathbf{R}_k|$ 代替,它由 $\Delta\mathbf{R}_k = [[-r_k(i,j), r_k(i,j)]]$ 区间元素的上限组成,即

$$|\Delta\mathbf{R}_k| = [r_k(i,j)], \quad r_k(i,j) \geq 0 \quad (10.13)$$

需要注意,这里 $|\Delta\mathbf{R}_k|$ 是普通矩阵(不是区间矩阵),所以当用普通矩阵求逆 $[\mathbf{C}_k \mathbf{M}_{k-1} \mathbf{C}_k^T + \Delta\mathbf{R}_k]^{-1}$ 来替换区间矩阵求逆 $[[\mathbf{C}_k^l] \mathbf{M}_{k-1}^l [\mathbf{C}_k^l]^T + \mathbf{R}_k]^{-1}$ 时,矩阵的逆变得更容易。更重要的是,当式(10.13)中的扰动矩阵 $\Delta\mathbf{C}_k = \mathbf{0}$ 时,意味着式(10.2)所示系统的量测方程和式(10.1)所示的标称系统模型一样是准确的,有 $|\Delta\mathbf{R}_k| = \mathbf{R}_k$ 。

因此,用 $|\Delta\mathbf{R}_k|$ 代替 $\Delta\mathbf{R}_k$,可以得到下面的次优区间卡尔曼滤波方案。

一种次优区间卡尔曼滤波算法

主过程

$$\begin{aligned}\hat{\mathbf{x}}_0^I &= \mathbf{E}(\mathbf{x}_0^I) \\ \hat{\mathbf{x}}_k^I &= \mathbf{A}_k^I \hat{\mathbf{x}}_{k-1}^I + \mathbf{G}_k^I [\mathbf{v}_k^I - \mathbf{C}_k^I \mathbf{A}_{k-1}^I \hat{\mathbf{x}}_{k-1}^I] \\ k &= 1, 2, \dots\end{aligned}\quad (10.14)$$

从过程

$$\begin{aligned}\mathbf{P}_0^I &= \text{Var}(\mathbf{x}_0^I) \\ \mathbf{M}_{k-1}^I &= \mathbf{A}_{k-1}^I \mathbf{P}_{k-1}^I [\mathbf{A}_{k-1}^I]^T + \mathbf{B}_{k-1}^I \mathbf{Q}_{k-1} [\mathbf{B}_{k-1}^I]^T \\ \mathbf{G}_k^I &= \mathbf{M}_{k-1}^I [\mathbf{C}_k^I]^T [\mathbf{C}_k^I \mathbf{M}_{k-1}^I \mathbf{C}_k^T + |\Delta\mathbf{R}_k|]^{-1} \\ \mathbf{P}_k^I &= [\mathbf{I} - \mathbf{G}_k^I \mathbf{C}_k^I] \mathbf{M}_{k-1}^I [\mathbf{I} - \mathbf{G}_k^I \mathbf{C}_k^I]^T + [\mathbf{G}_k^I] \mathbf{R}_k [\mathbf{G}_k^I]^T \\ k &= 1, 2, \dots\end{aligned}\quad (10.15)$$

最后,注意到式(10.13)给出的最坏情况下的矩阵 $|\Delta\mathbf{R}_k|$ 包含了最大的可能干扰。在某种意义上,是能够给出稳态数值逆的“最好的”矩阵。另一个可能的近似为,如果 $\Delta\mathbf{C}_k$ 很小,可以简单地使用 $|\Delta\mathbf{R}_k| \approx \mathbf{R}_k$ 。在一些像后面章节中将要讨论的雷达跟踪系统的特殊系统中,特殊的技术也可能用来改善次优区间滤波的速度或精度。

10.2.3 目标跟踪的例子

本节介绍一个计算机仿真的例子来比较区间卡尔曼滤波和标准卡尔曼滤波。对雷达跟踪系统(3.26)的简化,可以参考式(4.22)、式(5.22)、式(6.43):

$$\begin{cases} \mathbf{x}_{k+1} = \begin{bmatrix} 1 & h^I \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \xi_k \\ v_k = [1 \ 0] \mathbf{x}_k + \eta_k \end{cases} \quad (10.16)$$

式中的基本假设与系统(10.2)中介绍的一致。

假定系统有一个小区间不确定性:

$$h^I = [h - \Delta h, h + \Delta h] = [0.01 - 0.001, 0.01 + 0.001] = [0.009, 0.011]$$

式中建模误差 Δh 取为设定值 $h=0.01$ 的10%。

假设其他的给定数据为

$$\mathbf{E}(\mathbf{x}_0) = \begin{bmatrix} x_{01} \\ x_{02} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{Var}(\mathbf{x}_0) = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.0 \\ 0.0 & 0.5 \end{bmatrix}$$

$$\mathbf{Q}_k = \begin{bmatrix} q & 0 \\ 0 & q \end{bmatrix} = \begin{bmatrix} 0.1 & 0.0 \\ 0.0 & 0.1 \end{bmatrix}, \quad R_k = r = 0.1$$

对于该模型,应用区间卡尔曼滤波(10.11)~(10.12),有

$$\begin{aligned}\mathbf{M}_{k-1}^I &= \begin{bmatrix} h^I [2P_{k-1}^I(1,0) + h^I P_{k-1}^I(1,1)] + P_{k-1}^I(0,0) + q & P_{k-1}^I(0,1) + h^I P_{k-1}^I(1,1) \\ P_{k-1}^I(1,0) + h^I P_{k-1}^I(1,1) & P_{k-1}^I(1,1) + q \end{bmatrix} \\ &:= \begin{bmatrix} M_{k-1}^I(0,0) & M_{k-1}^I(0,1) \\ M_{k-1}^I(1,0) & M_{k-1}^I(1,1) \end{bmatrix} \\ \mathbf{G}_k^I &= \begin{bmatrix} 1 - r/(M_{00}^I + r) \\ M_{10}^I / (M_{00}^I + r) \end{bmatrix} := \begin{bmatrix} G_{k,1}^I \\ G_{k,2}^I \end{bmatrix} \\ \mathbf{P}_k^I &= \begin{bmatrix} rG_{k,1}^I & rG_{k,2}^I \\ rG_{k,2}^I & q + [P_{k-1}^I(1,1)[P_{k-1}^I(0,0) + q + r] - [P_{k-1}^I(0,1)]^2] / (M_{k-1}^I(0,0) + r) \end{bmatrix} \\ &:= \begin{bmatrix} P_k^I(0,0) & P_k^I(0,1) \\ P_k^I(1,0) & P_k^I(1,1) \end{bmatrix}\end{aligned}$$

在上面的式子中,矩阵 \mathbf{M}_{k-1}^I 和 \mathbf{P}_k^I 都是对称矩阵,所以, $M_{k-1}^I(0,1) = M_{k-1}^I(1,0)$, $P_{k-1}^I(0,1) = P_{k-1}^I(1,0)$, 根据滤波算法可以得到:

$$\begin{bmatrix} \hat{x}_{k,1}^I \\ \hat{x}_{k,2}^I \end{bmatrix} = \begin{bmatrix} [r(\hat{x}_{k-1,1}^I + h^I \hat{x}_{k-1,2}^I) + M_{k-1}^I(0,0)y_k] / M_{k-1}^I(0,0) \\ \hat{x}_{k-1,2}^I + G_{k,1}^I(y_k - \hat{x}_{k-1,1}^I - h^I \hat{x}_{k-1,2}^I) \end{bmatrix}$$

区间卡尔曼滤波和标准卡尔曼滤波的 $\hat{x}_{k,1}$ 仿真结果及比较见图 10.1 和图 10.2, 其中标准卡尔曼滤波用了确定值 h 。从这两幅图可以看出,算法(10.11)~(10.12)介绍的新方法得到的信号估计轨迹的上界和下界,包含了利用式(3.25)标准卡尔曼滤波算法得到的信号估计轨迹。这两条界线包含了区间系统(10.2)的所有可能的最优估计。

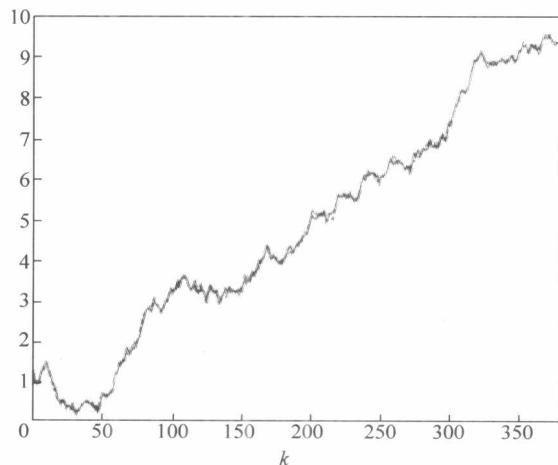


图 10.1 区间卡尔曼滤波估计结果(状态分量 1)

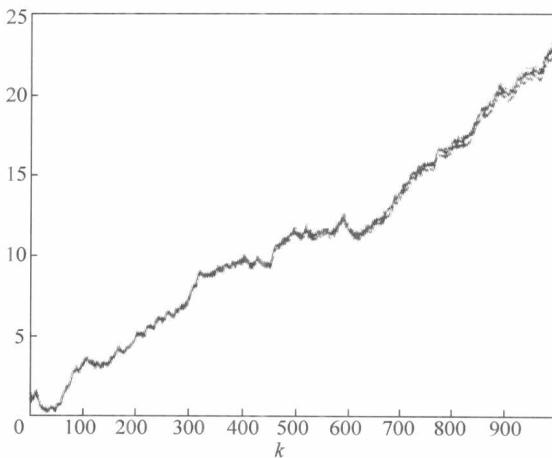


图 10.2 区间卡尔曼滤波估计结果(状态分量 2)

10.3 加权平均区间卡尔曼滤波

从图 10.1 和图 10.2 可以看出, 区间卡尔曼滤波能够产生标准卡尔曼滤波算法所有可能得到的最优轨迹的上界和下界。但是, 随着递推的不断进行, 上下界逐渐分开。这里应该再次强调, 这个看上去发散的结果不是因为滤波算法, 而是由于区间系统模型。也就是说, 即使模型没有噪声也没有进行滤波, 区间系统的上下界轨迹本身也会逐渐分开。所以这种现象是区间系统固有的, 虽然它本身是用来对不确定动态系统建模。

为了避免使用区间系统模型的这种发散, 一个可行方案是对两条边界包含的所有可能的最优估计轨迹进行加权平均。一个更简便的方法是简单地对两条边界估计进行加权平均。例如, 取图 10.1 和图 10.2 中两个区间滤波估计的某个加权平均, 可以分别得到图 10.3 和图 10.4。

最后, 非常重要的是, 该平均与分别对两个边界系统(10.9)和(10.10)采用两个标准卡尔曼滤波所得到的轨迹进行平均是有着本质上的不同的。主要是因为, 图 10.1 和图 10.2 所示的滤波轨迹的两个边界, 包含了所有可能的最优估计, 但是对两个边界系统应用标准卡尔曼滤波的估计, 并没有覆盖所有的解(正如前面指出的, 见习题 10.5)。

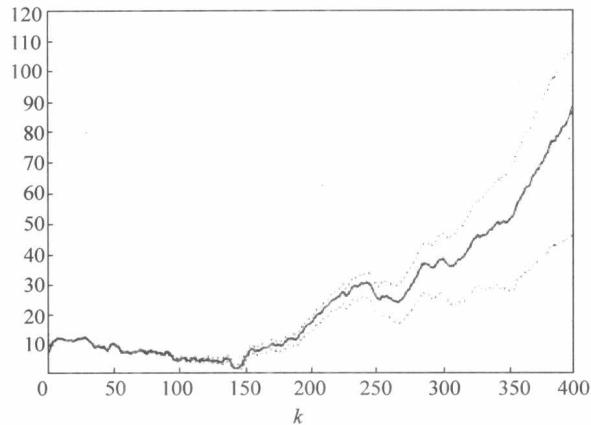


图 10.3 加权平均区间卡尔曼滤波(状态分量 1)

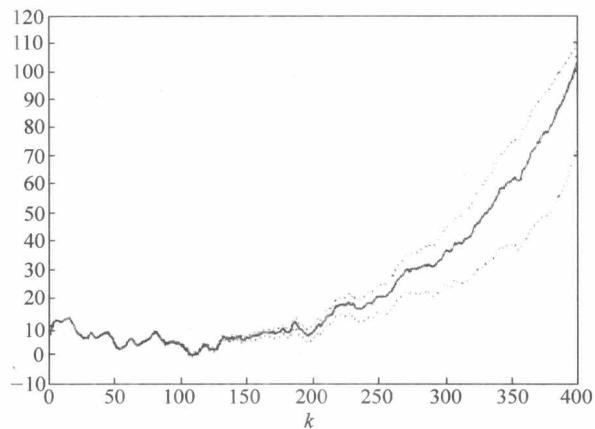


图 10.4 加权平均区间卡尔曼滤波(状态分量 2)

练习

10.1 对于三个区间 X 、 Y 和 Z , 证明:

$$X+Y=Y+X;$$

$$Z+(X+Y)=(Z+X)+Y;$$

$$XY=YX;$$

$$Z(XY) = (ZX)Y;$$

$X+0=0+X=X$ 和 $X0=0X=0$, 其中 $0=[0,0]$;

$XI=IX=X$, 其中 $I=[1,1]$;

$Z(X+Y) \subseteq ZX+ZY$, 其中“=”只在下面条件下成立:

- (a) $Z=[z,z]$;
- (b) $X=Y=0$;
- (c) 对所有 $x \in X$ 和 $y \in Y$, $xy \geq 0$.

10.2 设 \mathbf{A}, \mathbf{B} 和 \mathbf{C} 分别是适当维数的普通常值矩阵, $\mathbf{A}^I, \mathbf{B}^I$ 和 \mathbf{C}^I 分别是适当维数的区间矩阵。验证:

- (a) $\mathbf{A}^I \pm \mathbf{B}^I = \{\mathbf{A} \pm \mathbf{B} \mid \mathbf{A} \in \mathbf{A}^I, \mathbf{B} \in \mathbf{B}^I\}$;
- (b) $\mathbf{A}^I \mathbf{B} = \{\mathbf{AB} \mid \mathbf{A} \in \mathbf{A}^I\}$;
- (c) $\mathbf{A}^I + \mathbf{B}^I = \mathbf{B}^I + \mathbf{A}^I$;
- (d) $\mathbf{A}^I + (\mathbf{B}^I + \mathbf{C}^I) = (\mathbf{A}^I + \mathbf{B}^I) + \mathbf{C}^I$;
- (e) $\mathbf{A}^I + \mathbf{0} = \mathbf{0} + \mathbf{A}^I = \mathbf{A}^I$;
- (f) $\mathbf{A}^I \mathbf{I} = \mathbf{I} \mathbf{A}^I = \mathbf{A}^I$;
- (g) 分配率
 - (g. 1) $(\mathbf{A}^I + \mathbf{B}^I) \mathbf{C}^I \subseteq \mathbf{A}^I \mathbf{C}^I + \mathbf{B}^I \mathbf{C}^I$;
 - (g. 2) $\mathbf{C}^I (\mathbf{A}^I + \mathbf{B}^I) \subseteq \mathbf{C}^I \mathbf{A}^I + \mathbf{C}^I \mathbf{B}^I$;
- (h) $(\mathbf{A}^I + \mathbf{B}^I) \mathbf{C} = \mathbf{A}^I \mathbf{C} + \mathbf{B}^I \mathbf{C}$;
- (i) $\mathbf{C}(\mathbf{A}^I + \mathbf{B}^I) = \mathbf{CA}^I + \mathbf{CB}^I$;
- (j) 结合律
 - (j. 1) $\mathbf{A}^I (\mathbf{BC}) \subseteq (\mathbf{A}^I \mathbf{B}) \mathbf{C}$;
 - (j. 2) 如果 $\mathbf{C}^I = -\mathbf{C}^I$, 则 $(\mathbf{AB}^I) \mathbf{C}^I \subseteq \mathbf{A}(\mathbf{B}^I \mathbf{C}^I)$;
 - (j. 3) $\mathbf{A}(\mathbf{B}^I \mathbf{C}) = (\mathbf{AB}^I) \mathbf{C}$;
 - (j. 4) 如果 $\mathbf{B}^I = -\mathbf{B}^I$ 且 $\mathbf{C}^I = -\mathbf{C}^I$, 则 $\mathbf{A}^I (\mathbf{B}^I \mathbf{C}^I) = (\mathbf{A}^I \mathbf{B}^I) \mathbf{C}^I$ 。

10.3 证明定理 10.2。

10.4 证明式(10.7)和式(10.8)。

10.5 举一个一维的简单例子, 说明对区间系统(10.2)的两个边界系统(10.9)~(10.10)应用标准卡尔曼滤波, 得到的两个滤波轨迹不能包含区间系统的所有可能的最优估计解。

10.6 考虑跟踪一个具有不确定输入的弹道导弹的目标跟踪问题。这个物理问题可以用下面的简化的区间模型描述:

$$\begin{cases} \mathbf{x}_{k+1}^I = \mathbf{A}_k^I \mathbf{x}_k^I + \underline{\xi}_k \\ \mathbf{v}_k^I = \mathbf{C}_k^I \mathbf{x}_k^I + \underline{\eta}_k \end{cases}$$

式中 $\mathbf{x}^I = [x_1^I \ \cdots \ x_7^I]^T$ 。

$$\begin{aligned} A_{11} &= A_{12} = A_{13} = 1, & A_{44} &= -\frac{1}{2}gx_7 \frac{z^2 + x_4^2}{z} \\ A_{45} &= A_{54} = -\frac{1}{2}g \frac{x_7 x_4 x_5}{z}, & A_{46} &= A_{64} = -\frac{1}{2}g \frac{x_7 x_4 x_6}{z} \\ A_{47} &= -\frac{1}{2}gx_4 z, & A_{55} &= -\frac{1}{2}gx_7 \frac{z^2 + x_5^2}{z} \\ A_{56} &= A_{65} = -\frac{1}{2}g \frac{x_7 x_5 x_6}{z}, & A_{57} &= -\frac{1}{2}gx_5 z \\ A_{66} &= -\frac{1}{2}gx_7 \frac{z^2 + x_6^2}{z}, & A_{67} &= -\frac{1}{2}gx_6 z \\ A_{76} &= -K^I x_7, & A_{77} &= -K^I x_6 \end{aligned}$$

其他所有的 $A_{ij} = 0$, 其中 K^I 是不确定系统参数, g 是重力常数, $z = \sqrt{x_4^2 + x_5^2 + x_6^2}$, 又

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

系统噪声和量测噪声序列为相互独立的零均值高斯噪声, 分别有方差 $\{\mathbf{Q}_k\}$ 和 $\{\mathbf{R}_k\}$ 。用下面的数据, 对该模型应用区间卡尔曼滤波算法:

$$g = 0.981;$$

$$K^I = [2.3 \times 10^{-5}, 3.5 \times 10^{-5}];$$

$$\begin{aligned} \mathbf{x}_0^I &= [3.2 \times 10^5, 3.2 \times 10^5, 2.1 \times 10^5, -1.5 \times 10^4, -1.5 \times 10^4, \\ &\quad -8.1 \times 10^3, 5 \times 10^{-10}]^T; \end{aligned}$$

$$\mathbf{P}_0^I = \text{diag}\{10^6, 10^6, 10^6, 10^6, 10^6, 1.286 \times 10^{-13} \exp\{-23.616\}\};$$

$$\mathbf{Q}_k = \frac{1}{k+1} \text{diag}\{0, 0, 0, 100, 100, 100, 2.0 \times 10^{-18}\};$$

$$\mathbf{R}_k = \frac{1}{k+1} \text{diag}\{150, 150, 150\}.$$

第 11 章

小波卡尔曼滤波

除了前面章节讨论的卡尔曼滤波算法之外,还有很多其他可用的时域数字滤波计算方案。在这些方案中,最令人激动的可能就是小波算法,它可用于多尺度信号处理(估计或滤波)和信号多分辨率分析。本章将专门介绍小波卡尔曼滤波这一有效的技术,通过解决一个随机信号的同时估计和分解的具体应用例子,来说明采用小波的卡尔曼滤波器组的方案。

11.1 小波初步

小波的概念是 20 世纪 80 年代初被提出来的,通过两个简单的操作“平移”和“伸缩”,由一个“小波基”函数族来实现。设 $\psi(t)$ 是这样的一个小波基函数,结合伸缩常数 a 和平移常数 b ,可以得到一系列形式为 $\psi((t-b)/a)$ 的小波函数。以这一系列小波函数为积分核来定义一种积分变换,称为小波积分变换(IWT, integral wavelet transform) :

$$(W_\psi f)(b, a) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \overline{\psi((t-b)/a)} dt, \quad f \in L^2 \quad (11.1)$$

可以根据 a 和 b 的值在不同的位置和不同的尺度来分析函数(或信号)。注意小波基 $\psi(t)$ 起到了时间窗口的作用,其宽度随着公式(11.1)中尺度因子 a 的减小而变窄。因此,如果其傅里叶变换 $\hat{\psi}(\omega)$ 的频率 ω 和尺度因子 a 成反比,这样由 $\psi(t)$ 产生的时间窗口变窄,可以用来研究高频信号,或时间窗口变宽来对低频情况进行观测。此外,小波基 $\psi(t)$ 的傅里叶变换也是一个窗口函数。那么 IWT 的 $(W_\psi f)(b, a)$ 就可以在 $t=b$ 附近及尺度因子 a 所定义的频带宽度内用

来对 $f(t)$ 进行时域与频域定位和分析。

11.1.1 小波基础

“多分辨率分析”为研究小波提供了一个简洁方式。设 L^2 是连续时间域 $(-\infty, \infty)$ 内的实值能量有限函数的空间, 其内积定义为 $\langle f, g \rangle = \int_{-\infty}^{\infty} f(t) \bar{g}(t) dt$, 模定义为 $\|f\|_{L^2} = \sqrt{|\langle f, f \rangle|}$ 。如果存在一些窗口函数(见练习 11.1) $\phi(t) \in L^2$, 满足下列特性, 则 L^2 上闭子空间组成的区间套 $\{V_k\}$ 称为 L^2 的一个多分辨率分析:

(i) 对于每一个整数 k , 集合

$$\{\phi_{kj}(t) := \phi(2^k t - j) : j = \dots, -1, 0, 1, \dots\}$$

是 V_k 的 Riesz 基, 也就是说其线性张开在 V_k 中是稠密的, 并且对所有的 $\{c_j\} \in l^2$ 和每一个 k , 均有

$$\alpha \|\{c_j\}\|_{l^2}^2 \leq \left\| \sum_{j=-\infty}^{\infty} c_j \phi_{kj} \right\|_{L^2}^2 \leq \beta \|\{c_j\}\|_{l^2}^2 \quad (11.2)$$

式中 $\|\{c_j\}\|_{l^2} = \sqrt{\sum_{j=-\infty}^{\infty} |c_j|^2}$;

(ii) V_k 的组合在 L^2 中是稠密的;

(iii) 所有 V_k 的交集是零函数;

(iv) 当且仅当 $f(2t) \in V_{k+1}$ 时, $f(t) \in V_k$ 。

设 V_{k+1} 为 V_k 及其正交补 W_k 的直和, 表示为

$$V_{k+1} = V_k \oplus W_k \quad (11.3)$$

则对于所有的 $k \neq n$, 有 $W_k \perp W_n$, 并且整个 L^2 空间是子空间 W_k 的正交和, 即

$$L^2 = \bigoplus_{k=-\infty}^{\infty} W_k \quad (11.4)$$

假设存在一个函数 $\psi(t) \in W_0$, 使得每个 $\psi(t)$ 及其傅里叶变换 $\hat{\psi}(\omega)$ 在 $\pm\infty$ 处都能足够快的衰减(见练习 11.2), 并且对于任一整数 k ,

$$\{\psi_{kj}(t) := 2^{k/2} \psi(2^k t - j) : j = \dots, -1, 0, 1, \dots\} \quad (11.5)$$

是 W_k 的一个标准正交基, 则 $\psi(t)$ 称为是一个小波(见练习 11.3)。

设 $\tilde{\psi}(t) \in W_0$ 是 $\psi(t)$ 的对偶, 即

$$\int_{-\infty}^{\infty} \tilde{\psi}(t-i) \overline{\psi(t-j)} dt = \delta_{ij} \quad i, j = \dots, -1, 0, 1, \dots$$

则 $\tilde{\psi}(t)$ 和 $\hat{\psi}(\omega)$ 在时间域和频率域都分别是窗口函数。如果 $\tilde{\psi}(t)$ 在 IWT 分析中用作小波基, 可以得到实时算法, 由 $b=j/2^k$ 时刻和 $a=2^{-k}$ 尺度定义的第 k 个频

率带来确定 $(W_{\tilde{\psi}} f)(b, a)$ 。同样, $f(t)$ 也可以根据这些位置的 $(W_{\tilde{\psi}} f)(b, a)$ 的信息来实时重构。

更加精确地,由定义

$$\psi_{kj}(t) = 2^{k/2} \psi(2^k t - j) \quad (11.6)$$

任何函数 $f(t) \in L^2$ 都可以表示为小波分量的线性线合

$$f(t) = \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} d_j^k \psi_{k,j}(t) \quad (11.7)$$

式中

$$d_j^k = (W_{\tilde{\psi}} f)(j2^{-k}, 2^{-k}) \quad (11.8)$$

由式(11.3),在 l^2 中存在两个序列 $\{a_j\}$ 和 $\{b_j\}$,使得对所有的整数 l ,

$$\phi(2t - l) = \sum_{j=-\infty}^{\infty} [a_{l-2j} \phi(t - j) + b_{l-2j} \psi(t - j)] \quad (11.9)$$

从特性(i)和式(11.3)可以看出, l^2 中存在唯一确定的两个序列 $\{p_j\}$ 和 $\{q_j\}$, 满足

$$\phi(t) = \sum_{j=-\infty}^{\infty} p_j \phi(2t - j) \quad (11.10)$$

$$\psi(t) = \sum_{j=-\infty}^{\infty} q_j \phi(2t - j) \quad (11.11)$$

由序列对 $(\{a_j\}, \{b_j\})$ 得到小波变换 $\{d_j^k\}$ 的金字塔分解算法,而由序列对 $(\{p_j\}, \{q_j\})$ 得到小波变换 $\{d_j^k\}$,组成重新构建 $f(t)$ 的金字塔合成算法。

注意到,如果 $\phi(t)$ 选择为一个 B 样条函数,则可以得到一个紧支撑的小波 $\psi(t)$,其对偶值 $\tilde{\psi}(t)$ 在 $\pm\infty$ 处指数衰减,并且 $\tilde{\psi}(t)$ 的 IWT 具有线性相位。进一步地, $\{a_j\}$ 和 $\{b_j\}$ 具有较快的指数衰减,同时序列 $\{p_j\}$ 和 $\{q_j\}$ 是有限的。需要提及的是,样条小波 $\psi(t)$ 和 $\tilde{\psi}(t)$ 有显式表示,并且可以很容易实现。

11.1.2 离散小波变换和滤波器组

对于给定的确定标量信号 $\{x(i, n)\} \in l^2$,在固定的尺度 i ,通过脉冲响应为 $\{h(n)\}$ 的半宽低通滤波器,可以得到较低分辨率的信号。或者说,低分辨率信号序列(用 L 表示)可以通过对低通滤波器的输出下采样获得,即 $h(n) \rightarrow h(2n)$ 。因此,

$$x_L(i-1, n) = \sum_{k=-\infty}^{\infty} h(2n-k) x(i, k) \quad (11.12)$$

式(11.12)定义了一个从 l^2 到其自身的映射。作为 $x_L(i-1, n)$ 的补充,小波系

数为 $\{x_H(i-1,n)\}$,可以通过一个脉冲响应为 $\{g(n)\}$ 的高通滤波器,再对高通滤波的输出进行下采样计算得到。这样就有

$$x_H(i-1,n) = \sum_{k=-\infty}^{\infty} g(2n-k)x(i,k) \quad (11.13)$$

原始信号 $\{x(i,n)\}$ 可以由滤波和采样(低分辨率)信号 $\{x_L(i-1,n)\}$ 和 $\{x_H(i-1,n)\}$ 重构。为了生成完好的重构信号,滤波器 $\{h(n)\}$ 和 $\{g(n)\}$ 必须满足一些约束条件。最重要的约束条件就是滤波器脉冲响应组成正交集。因此,式(11.12)和式(11.13)可以被看作原始信号在正交基上的分解,而重构可以认为是正交投影的和,见下式:

$$x(i,n) = \sum_{k=-\infty}^{\infty} h(2k-n)x_L(i-1,k) + \sum_{k=-\infty}^{\infty} g(2k-n)x_H(i-1,k) \quad (11.14)$$

式(11.12)和式(11.13)定义的操作称为离散(前向)小波变换,而离散小波逆变换可以由式(11.14)来定义。

为了能对 $\{h(n)\}$ 和 $\{g(n)\}$ 都使用FIR(finite impulse response)滤波器(也就是说,这两个序列有有限长度, L),规定

$$g(n) = (-1)^n h(L-1-n) \quad (11.15)$$

式中 L 必须是满足这一关系的偶数。

显然,一旦低通滤波器 $\{h(n)\}$ 确定了,高通滤波器也就确定了。

离散小波变换可以通过倍频滤波器组来实现,如图11.1(b)所示,其中只描述了三层变换。在不同层不同分解的尺度见图11.1(a)。

对于有限长度的确定信号序列,用算子形式来描述小波变换更方便。考虑序列信号在第 i 层长度为 M 的分解:

$$\underline{\mathbf{X}}_k^i = [x(i,k-M+1), x(i,k-M+2), \dots, x(i,k)]^T$$

式(11.12)和式(11.13)可以写为以下的算子形式:

$$\underline{\mathbf{X}}_{k_L}^{i-1} = \mathbf{H}^{i-1} \underline{\mathbf{X}}_k^i, \quad \underline{\mathbf{X}}_{k_H}^{i-1} = \mathbf{G}^{i-1} \underline{\mathbf{X}}_k^i$$

式中,算子 \mathbf{H}^{i-1} 和 \mathbf{G}^{i-1} 为低通和高通滤波器(式(11.12)和式(11.13)中的 $\{h(n)\}$ 和 $\{g(n)\}$),第 i 层映射到第 $i-1$ 层的响应。同理,从第 $i-1$ 层映射到第 i 层,式(11.14)的算子形式可以写为(见练习11.4)

$$\underline{\mathbf{X}}_k^i = (\mathbf{H}^{i-1})^T \underline{\mathbf{X}}_{k_L}^{i-1} + (\mathbf{G}^{i-1})^T \underline{\mathbf{X}}_{k_H}^{i-1} \quad (11.16)$$

另一方面,正交约束也可以用算子的形式表示:

$$(\mathbf{H}^{i-1})^T \mathbf{H}^{i-1} + (\mathbf{G}^{i-1})^T \mathbf{G}^{i-1} = \mathbf{I}$$

$$\begin{bmatrix} \mathbf{H}^{i-1} (\mathbf{H}^{i-1})^T & \mathbf{H}^{i-1} (\mathbf{G}^{i-1})^T \\ \mathbf{G}^{i-1} (\mathbf{H}^{i-1})^T & \mathbf{G}^{i-1} (\mathbf{G}^{i-1})^T \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{bmatrix}$$

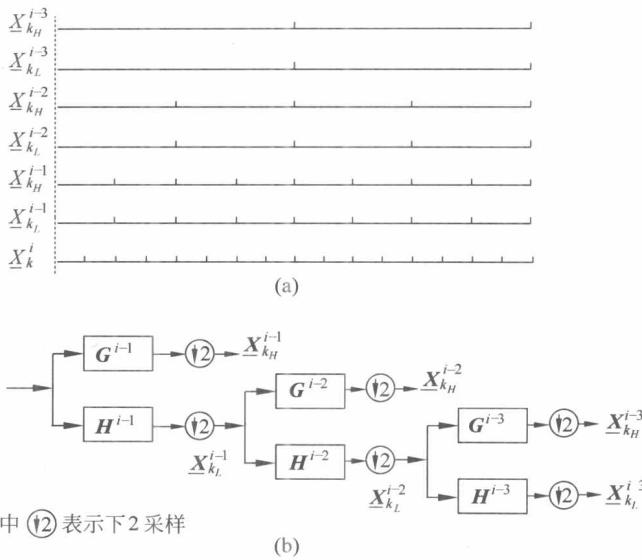


图 11.1 离散小波变换
(a) 信号分解; (b) 两通道滤波器组

信号的多层次同时分解可以通过滤波器组来实现。例如,将 $\underline{\mathbf{X}}_k^i$ 分解为三层,如图 11.1 所示,可以应用下面的综合变换:

$$\begin{bmatrix} \underline{\mathbf{X}}_{k_L}^{i-3} \\ \underline{\mathbf{X}}_{k_H}^{i-3} \\ \underline{\mathbf{X}}_{k_H}^{i-2} \\ \underline{\mathbf{X}}_{k_L}^{i-2} \\ \underline{\mathbf{X}}_{k_H}^{i-1} \end{bmatrix} = \mathbf{T}^{i-3|i} \underline{\mathbf{X}}_k^i$$

式中

$$\mathbf{T}^{i-3|i} = \begin{bmatrix} \mathbf{H}^{i-3} \mathbf{H}^{i-2} \mathbf{H}^{i-1} \\ \mathbf{G}^{i-3} \mathbf{H}^{i-2} \mathbf{H}^{i-1} \\ \mathbf{G}^{i-2} \mathbf{H}^{i-1} \\ \mathbf{G}^{i-1} \end{bmatrix}$$

是一个正交矩阵,同时将 $\underline{\mathbf{X}}_k^i$ 映射到滤波器组的三层。

11.2 信号估计和分解

在随机信号估计和分解中,通用的方法是先用量测数据估计未知信号,然后根据分辨率需求来分解所估计的信号。这两步通常是离线的,在实时应用中是

不可取的。

本节将介绍随机信号的同时最优估计和多分辨率分解技术,用来作为小波卡尔曼滤波的一个例子。基于离散小波变换,并用卡尔曼滤波器组推导出同时完成估计和分解的算法。该算法保持了卡尔曼滤波方案在估计时的优点,得到了未知信号的最优(线性、无偏、最小误差方差)估计,以递推的方式应用从含噪声信号中采样获得的数据。

本节所提出的方案具有以下特性:首先,一步就确定了被估计信号,而不是原来的两步方法,并且得到的信号具有所希望的分解;其次,该算法中采用了递推卡尔曼滤波方案,以便同时得到未知并且含噪声信号的最优估计和分解;最后,整个信号处理过程是在线的,即是实时的,对输入一批新的量测数据,能以需要的分解形式输出信号的估计。在这个过程中,信号首先被分解为块,然后对数据块进行滤波。该方案基于当前的量测数据和之前的最优估计值,就可以得到当前的估计值。这里,数据块的长度由需要分解的层数来决定。倍频滤波器组被用作多分辨分析的有效工具。

11.2.1 随机信号的估计和分解

考虑一个一维随机信号序列 $\{x(N, k)\}$ 的最高分辨率层(N 层),它符合下式:

$$x(N, k+1) = A(N, k)x(N, k) + \xi(N, k) \quad (11.17)$$

量测为

$$v(N, k) = C(N, k)x(N, k) + \eta(N, k) \quad (11.18)$$

式中 $\{\xi(N, k)\}$ 和 $\{\eta(N, k)\}$ 是相互独立的零均值高斯噪声序列,方差分别为 $Q(N, k)$ 和 $R(N, k)$ 。

给定一个量测序列 $\{v(N, k)\}$,估计和分解随机信号 $\{X(N, k)\}$ 的常规方法是按次序进行:首先,在最高分辨率水平找到一个估计 $\hat{x}(N, k)$;其次,应用小波变换将其分解到不同的分辨率上去。

下面将要介绍同时完成随机信号的估计和分解任务的算法。为了简便,仅讨论两层分解和估计,例如,从 N 层到 $N-1$ 和 $N-2$ 层。而在其他所有层,同时进行估计和分解的过程是完全一样的。

选择长度为 $M=2^2=4$ 的数据块,选择基2是为了便于设计一个倍频滤波器组,指数2为小波分解的层数。在 k 时刻,有

$$\underline{\mathbf{X}}_k^N = [x(N, k-3), x(N, k-2), x(N, k-1), x(N, k)]^\top$$

其数据块形式的等效动态系统进一步讨论如下。

为了理论上的方便,系统方程和量测方程式(11.17)和式(11.18)假设是时

不变的，并在分解层 N 上进行讨论。以长度 M 为间隔实现该过程，则

$$x(N, k+1) = Ax(N, k) + \xi(N, k) \quad (11.19)$$

或

$$x(N, k+1) = A^2x(N, k-1) + A\xi(N, k-1) + \xi(N, k) \quad (11.20)$$

或

$$\begin{aligned} x(N, k+1) &= A^3x(N, k-2) + A^2\xi(N, k-2) + A\xi(N, k-1) + \xi(N, k) \\ &\quad (11.21) \end{aligned}$$

或

$$\begin{aligned} x(N, k+1) &= A^4x(N, k-3) + A^3\xi(N, k-3) + A^2\xi(N, k-2) \\ &\quad + A\xi(N, k-1) + \xi(N, k) \quad (11.22) \end{aligned}$$

取式(11.19)、(11.20)、(11.21)和(11.22)的平均，可得

$$\begin{aligned} x(N, k+1) &= \frac{1}{4}A^4x(N, k-3) + \frac{1}{4}A^3x(N, k-2) \\ &\quad + \frac{1}{4}A^2x(N, k-1) + \frac{1}{4}Ax(N, k) + \xi(1) \end{aligned}$$

式中

$$\xi(1) = \frac{1}{4}A^3\xi(N, k-3) + \frac{1}{2}A^2\xi(N, k-2) + \frac{3}{4}A\xi(N, k-1) + \xi(N, k) \quad (11.23)$$

同理，所有的过程都采用上述相同的方法，最后可以得到下面的一个数据块形式的动态系统（见练习 11.5）：

$$\underbrace{\begin{bmatrix} x(N, k+1) \\ x(N, k+2) \\ x(N, k+3) \\ x(N, k+4) \end{bmatrix}}_{\underline{x}_{k+1}^N} = \underbrace{\begin{bmatrix} \frac{1}{4}A^4 & \frac{1}{4}A^3 & \frac{1}{4}A^2 & \frac{1}{4}A \\ 0 & \frac{1}{3}A^4 & \frac{1}{3}A^3 & \frac{1}{3}A^2 \\ 0 & 0 & \frac{1}{2}A^4 & \frac{1}{2}A^3 \\ 0 & 0 & 0 & A^4 \end{bmatrix}}_{\bar{A}} \times \underbrace{\begin{bmatrix} x(N, k-3) \\ x(N, k-2) \\ x(N, k-1) \\ x(N, k) \end{bmatrix}}_{\underline{x}_k^N} + \underbrace{\begin{bmatrix} \xi(1) \\ \xi(2) \\ \xi(3) \\ \xi(4) \end{bmatrix}}_{\bar{Q}^N} \quad (11.24)$$

式中 $\xi(i), i=2, 3, 4$ ，具有相似的定义，且

$$E\{\bar{W}_k^N\} = \mathbf{0}, \quad E\{\bar{W}_k^N (\bar{W}_k^N)^T\} = \bar{Q}$$

其中 \bar{Q} 的元素由下式给出：

$$\bar{q}_{11} = \frac{1}{16}A^6Q + \frac{1}{4}A^4Q + \frac{9}{16}A^2Q + Q, \quad \bar{q}_{12} = \frac{1}{6}A^5Q + \frac{1}{2}A^3Q + AQ$$

$$\begin{aligned}
 \bar{q}_{13} &= \frac{3}{8} A^4 Q + A^2 Q, & \bar{q}_{14} &= A^3 Q, & \bar{q}_{21} &= \bar{q}_{12} \\
 \bar{q}_{22} &= \frac{1}{9} A^6 Q + \frac{4}{9} A^4 Q + A^2 Q + Q, & \bar{q}_{23} &= \frac{1}{3} A^5 Q + A^3 Q + A Q \\
 \bar{q}_{24} &= A^4 Q + A^2 Q, & \bar{q}_{31} &= \bar{q}_{13}, & \bar{q}_{32} &= \bar{q}_{23} \\
 \bar{q}_{33} &= \frac{1}{4} A^6 Q + A^4 Q + A^2 Q + Q, & \bar{q}_{34} &= A^5 Q + A^3 Q + A Q \\
 \bar{q}_{41} &= \bar{q}_{14}, & \bar{q}_{42} &= \bar{q}_{24}, & \bar{q}_{43} &= \bar{q}_{34}, & \bar{q}_{44} &= A^6 Q + A^4 Q + A^2 Q + Q
 \end{aligned}$$

与式(11.24)相关的量测方程可以很容易的得出^{*}

$$\underbrace{\begin{bmatrix} v(N, k-3) \\ v(N, k-2) \\ v(N, k-1) \\ v(N, k) \end{bmatrix}}_{\underline{V}_k^N} = \underbrace{\begin{bmatrix} C & 0 & 0 & 0 \\ 0 & C & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & C \end{bmatrix}}_{\bar{C}} \underbrace{\begin{bmatrix} x(N, k-3) \\ x(N, k-2) \\ x(N, k-1) \\ x(N, k) \end{bmatrix}}_{\underline{X}_k^N} + \underbrace{\begin{bmatrix} \eta(N, k-3) \\ \eta(N, k-2) \\ \eta(N, k-1) \\ \eta(N, k) \end{bmatrix}}_{\underline{\eta}_k^N}$$
(11.25)

式中

$$E\{\underline{\Pi}_k^N\} = \mathbf{0}, \quad E\{\underline{\Pi}_k^N (\underline{\Pi}_k^N)^T\} = \mathbf{R} = \text{diag}\{R, R, R, R\}$$

采用两层分解, 可以得到

$$\begin{bmatrix} \underline{X}_{k_L}^{N-2} \\ \underline{X}_{k_H}^{N-2} \\ \underline{X}_{k_H}^{N-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}^{N-2} \mathbf{H}^{N-1} \\ \mathbf{G}^{N-2} \mathbf{H}^{N-1} \\ \mathbf{G}^{N-1} \end{bmatrix} \underline{X}_k^N = \mathbf{T}^{N-2|N} \underline{X}_k^N$$
(11.26)

将式(11.26)代入式(11.24), 可得

$$\begin{bmatrix} \underline{X}_{k+1L}^{N-2} \\ \underline{X}_{k+1H}^{N-2} \\ \underline{X}_{k+1H}^{N-1} \end{bmatrix} = \mathbf{A} \begin{bmatrix} \underline{X}_{k_L}^{N-2} \\ \underline{X}_{k_H}^{N-2} \\ \underline{X}_{k_H}^{N-1} \end{bmatrix} + \underline{W}_k^N$$
(11.27)

式中

$$\underline{W}_k^N = \mathbf{T}^{N-2|N} \bar{\underline{W}}_k^N, \quad E\{\underline{W}_k^N\} = \mathbf{0}, \quad E\{\underline{W}_k^N (\underline{W}_k^N)^T\} = \mathbf{Q}$$

$$\mathbf{A} = \mathbf{T}^{N-2|N} \bar{\mathbf{A}} (\mathbf{T}^{N-2|N})^T, \quad \mathbf{Q} = \mathbf{T}^{N-2|N} \bar{\mathbf{Q}} (\mathbf{T}^{N-2|N})^T$$

式(11.27)描述了一个可分解的动态系统。其相关的量测方程同样可以通过把式(11.26)代入式(11.25)得到, 有

$$\underline{V}_k^N = \mathbf{C} \begin{bmatrix} \underline{X}_{k_L}^{N-2} \\ \underline{X}_{k_H}^{N-2} \\ \underline{X}_{k_H}^{N-1} \end{bmatrix} + \underline{\Pi}_k^N$$
(11.28)

* 译者注: 此处式号原著似有误, 译者适当调整。

式中

$$\mathbf{C} = \bar{\mathbf{C}} (\mathbf{T}^{N-2|N})^T$$

如此得到了系统方程和量测方程式(11.27)和式(11.28)的各个分解量。下一个任务是利用量测数据估计这些量。对式(11.27)和式(11.28)应用卡尔曼滤波,可以得到这些分解量的最优估计。

11.2.2 一个随机游走的例子

本小节将研究一个一维有色噪声过程(也叫做布朗随机游走)。该随机过程可以表示为

$$x(N, k+1) = x(N, k) + \xi(N, k) \quad (11.29)$$

量测为

$$v(N, k) = x(N, k) + \eta(N, k) \quad (11.30)$$

式中 $\{\xi(N, k)\}$ 、 $\{\eta(N, k)\}$ 是相互独立的零均值高斯噪声序列, 方差分别是 $Q(N, k) = 0.1$ 、 $R(N, k) = 1.0$ 。

真实值 $\{x(N, k)\}$ 和量测值 $\{v(N, k)\}$ 的最高分辨率如图 11.2(a) 和 (b) 所示。采用模型(11.27)、(11.28), 在滤波器组应用 Haar 小波, 对前面给出的过程应用标准卡尔曼滤波, 可得到两层估计: $\hat{x}_{k_L}^{N-2}$ 、 $\hat{x}_{k_H}^{N-2}$ 和 $\hat{x}_{k_H}^{N-1}$ 。在线计算结果见图 11.2(c)(d) 和图 11.3(a)。

初看时, $\hat{x}_{k_H}^{N-2}$ 和 $\hat{x}_{k_H}^{N-1}$ 都像某种噪声。实际上, $\hat{x}_{k_H}^{N-2}$ 和 $\hat{x}_{k_H}^{N-1}$ 为真实信号高频部分的估计。可以通过把这些高频部分“加”到低频部分来构成多层估计。例如, $\hat{x}_{k_L}^{N-1}$ 是按下式, 通过组合 $\hat{x}_{k_L}^{N-2}$ 和 $\hat{x}_{k_H}^{N-2}$ 重构而成的:

$$\hat{x}_{k_L}^{N-1} = [(\mathbf{H}^{N-2})^T (\mathbf{G}^{N-2})^T] \begin{bmatrix} \hat{x}_{k_L}^{N-2} \\ \hat{x}_{k_H}^{N-2} \end{bmatrix} \quad (11.31)$$

见图 11.3(b)。

同理, $\hat{x}_{k_L}^N$ (在最高分辨率水平) 可以通过组合 $\hat{x}_{k_L}^{N-1}$ 和 $\hat{x}_{k_H}^{N-1}$ 来获得, 见图 11.3(c)。

为了比较性能, 标准卡尔曼滤波被直接应用到系统(11.29)和(11.30), 结果见图 11.3(d)。可以看出通过组合分解量的估计和通过卡尔曼滤波得到的在最高分辨率估计非常相似。为了在数量上比较这两种方法, 进行 200 次的蒙特卡罗仿真。这时估计和分解算法的平方根误差为 0.2940, 而标准卡尔曼滤波为 0.3104。这表明即使只有两层分解和估计, 前者比直接卡尔曼滤波效果好。如果允许分解的层数增加, 两者的差别将更加显著。

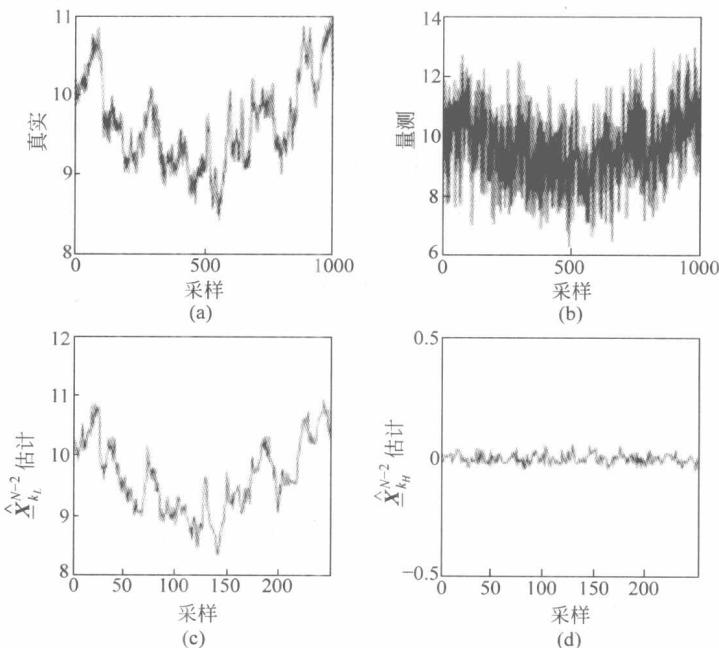


图 11.2 信号的小波分解

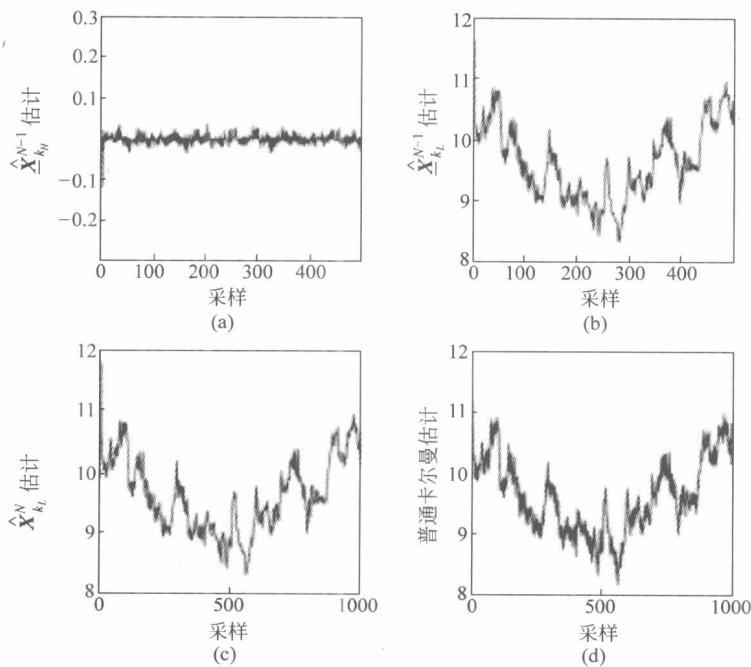


图 11.3 信号估计结果

练习

11.1 下面的 Haar 和三角函数是典型的时域窗函数：

$$\phi_H(t) = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{其他} \end{cases}$$

$$\phi_T(t) = \begin{cases} t, & 0 \leq t < 1 \\ 2-t, & 1 \leq t < 2 \\ 0, & \text{其他} \end{cases}$$

观察这两个函数，然后证明：

$$\phi_T(t) = \int_{-\infty}^{\infty} \phi_H(\tau) \phi_H(t - \tau) d\tau$$

其中 $\phi_H(t)$ 和 $\phi_T(t)$ 分别称为 0 度 B 样条和 1 度 B 样条。 n 度 B 样条由下式生成：

$$\phi_n(t) = \int_0^1 \phi_{n-1}(t - \tau) d\tau, \quad n = 2, 3, \dots$$

计算并描述 $\phi_2(t)$ 和 $\phi_3(t)$ （注意 $\phi_0 = \phi_H$ 和 $\phi_1 = \phi_T$ ）。

11.2 给出上面定义的 $\phi_H(t)$ 和 $\phi_T(t)$ 的傅里叶变换。

11.3 根据 $\phi_H(t)$ 和 $\phi_T(t)$ 的图示，画出 $\phi_{1,0}(t) = \phi_H(2t)$ 、 $\phi_{1,1}(t) = \phi_H(2t-1)$ 和 $\Phi(t) = \phi_{1,0}(t) + \phi_{1,1}(t)$ 。更进一步画出小波：

$$\Psi_H(t) = \phi_H(2t) - \phi_H(2t-1)$$

$$\Psi_T(t) = -\frac{1}{2}\phi_T(2t) - \frac{1}{2}\phi_T(2t-2) + \phi_T(2t-1)$$

11.4 证明式(11.16)。

11.5 证明式(11.24)和式(11.26)。

第 12 章

附录

这本教科书的目的是给出卡尔曼滤波理论基本但严密的介绍,同时简单涉及它的一些实时应用。全书没有期望覆盖该理论的所有基本原理,并且只包含了很有限的一些应用。在参考文献中有介绍不同应用的很多著作,包括 Anderson 和 Moore (1979), Balakrishnan (1984, 1987), Brammer 和 Sifflin (1989), Catlin (1989), Chen (1985), Chen, Chen and Hsu (1995), Goodwin 和 Sin (1984), Haykin (1986), Lewis (1986), Mendel (1987), Ruyymgaart 和 Soong (1985, 1989), Sorenson (1985), Stengel (1986), 和 Young (1984)。遗憾的是,在我们的详细介绍过程中,不得不忽略许多重要的主题。这一章将简要地介绍其中的一些主题,感兴趣的读者可以根据本书提供的参考文献进一步学习和研究。

12.1 卡尔曼平滑器

假设数据信息是在某一离散时间区间 $\{1, 2, \dots, N\}$ 上获得的。对于任意满足 $1 \leq K < N$ 的 K , 使用该时间区间上的所有数据信息(例如, 过去的、现在的和将来所有信息都使用)的状态向量 x_K 的最优估计 $\hat{x}_{K|N}$, 称为是 x_K 的一个(数字)平滑估计(对照第 2 章定义 2.1)。虽然该平滑问题稍微不同于本书中考虑的实时估计问题,但是仍有许多实际的应用。一个典型的应用就是对卫星轨道的确定,其中卫星轨道的估计允许在某个确定的时间周期之后进行。更精确地,考虑下面的具有固定终点时间 N 的线性确定/随机系统:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \boldsymbol{\Gamma}_k \boldsymbol{\xi}_k \\ \mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k + \boldsymbol{\eta}_k \end{cases}$$

式中 $1 \leq k < N$ 。这里假设对于所有 k , $\boldsymbol{\xi}_k$ 的方差矩阵 \mathbf{Q}_k 都是正定的(对照第 2 章和第 3 章,其中 \mathbf{Q}_k 只假设为非负定)。

卡尔曼滤波过程被用来寻找最优平滑估计 $\hat{\mathbf{x}}_{K|N}$, 其中“最优”仍是最小方差意义下的。首先, 定义利用数据信息 $\{\mathbf{v}_1, \dots, \mathbf{v}_K\}$ 的常规卡尔曼滤波估计 $\hat{\mathbf{x}}_{K|K}$ 为 $\hat{\mathbf{x}}_K^f$, 上标 f 表示该估计是一个前向(forward)过程。类似的, 定义 $\hat{\mathbf{x}}_K^b = \hat{\mathbf{x}}_{K|K+1}$ 为利用数据信息 $\{\mathbf{v}_{K+1}, \dots, \mathbf{v}_N\}$ 估计状态向量 \mathbf{x}_K 时所获得的后向(backward)最优预测。然后, 通过利用时间区间 $\{1, 2, \dots, N\}$ 上的所有数据信息得到最优平滑估计 $\hat{\mathbf{x}}_{K|N}$ (常称为卡尔曼平滑估计)(更详细的介绍见 Lewis (1986) 和 Balakrishnan(1984, 1987)):

$$\begin{cases} \mathbf{G}_K = \mathbf{P}_K^f \mathbf{P}_K^b (\mathbf{I} + \mathbf{P}_K^f \mathbf{P}_K^b)^{-1} \\ \mathbf{P}_K = (\mathbf{I} - \mathbf{G}_K) \mathbf{P}_K^f \\ \hat{\mathbf{x}}_{K|N} = (\mathbf{I} - \mathbf{G}_K) \hat{\mathbf{x}}_K^f + \mathbf{P}_K \hat{\mathbf{x}}_K^b \end{cases}$$

式中, \mathbf{P}_K^f 、 $\hat{\mathbf{x}}_K^f$ 和 \mathbf{P}_K^b 、 $\hat{\mathbf{x}}_K^b$ 分别利用下面的方法递推计算:

$$\begin{cases} \mathbf{P}_0^f = \text{Var}(\mathbf{x}_0) \\ \mathbf{P}_{k,k-1}^f = \mathbf{A}_{k-1} \mathbf{P}_{k-1}^f \mathbf{A}_{k-1}^\top + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^\top \\ \mathbf{G}_k^f = \mathbf{P}_{k,k-1}^f \mathbf{C}_k^\top (\mathbf{C}_k \mathbf{P}_{k,k-1}^f \mathbf{C}_k^\top + \mathbf{R}_k)^{-1} \\ \mathbf{P}_k^f = (\mathbf{I} - \mathbf{G}_k^f \mathbf{C}_k) \mathbf{P}_{k,k-1}^f \\ \hat{\mathbf{x}}_0^f = E(\mathbf{x}_0) \\ \hat{\mathbf{x}}_k^f = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1}^f + \mathbf{B}_{k-1} \mathbf{u}_{k-1} + \mathbf{G}_k^f (\mathbf{v}_k - \mathbf{C}_k \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1}^f - \mathbf{D}_k \mathbf{u}_{k-1}) \\ k = 1, 2, \dots, K \end{cases}$$

及

$$\begin{cases} \mathbf{P}_N^b = \mathbf{0} \\ \mathbf{P}_{k+1,N}^b = \mathbf{P}_{k+1}^b + \mathbf{C}_{k+1}^\top \mathbf{R}_{k+1}^{-1} \mathbf{C}_{k+1} \\ \mathbf{G}_k^b = \mathbf{P}_{k+1,N}^b \boldsymbol{\Gamma}_{k+1}^\top (\boldsymbol{\Gamma}_{k+1} + \mathbf{P}_{k+1,N}^b \boldsymbol{\Gamma}_{k+1}^\top + \mathbf{Q}_{k+1}^{-1})^{-1} \\ \mathbf{P}_k^b = \mathbf{A}_{k+1}^\top (\mathbf{I} - \mathbf{G}_k^b \boldsymbol{\Gamma}_{k+1}^\top) \mathbf{P}_{k,N}^b \mathbf{A}_{k+1} \\ \hat{\mathbf{x}}_N^b = \mathbf{0} \\ \hat{\mathbf{x}}_k^b = \mathbf{A}_{k+1}^\top (\mathbf{I} - \mathbf{G}_k^b \boldsymbol{\Gamma}_{k+1}^\top) \{ \hat{\mathbf{x}}_{k+1}^b + \mathbf{C}_{k+1}^\top \mathbf{R}_{k+1}^{-1} \mathbf{v}_{k+1} - (\mathbf{C}_{k+1}^\top \mathbf{R}_{k+1}^{-1} \mathbf{D}_{k+1} + \mathbf{P}_{k+1,N}^b \mathbf{B}_{k+1}) \mathbf{u}_{k+1} \} \\ k = N-1, N-2, \dots, K \end{cases}$$

12.2 α - β - γ - θ 跟踪器

考虑一个用状态空间表示的、以有色噪声为输入的线性时不变随机系统：

$$\begin{cases} \dot{\mathbf{x}}_{k+1} = \mathbf{A}\mathbf{x}_k + \boldsymbol{\Gamma}\underline{\boldsymbol{\xi}}_k \\ \mathbf{v}_k = \mathbf{C}\mathbf{x}_k + \underline{\boldsymbol{\eta}}_k \end{cases}$$

式中 $\mathbf{A}, \mathbf{C}, \boldsymbol{\Gamma}$ 分别是 $n \times n, q \times n, n \times p$ 阶常值矩阵, $1 \leq p, q \leq n$, 且

$$\begin{cases} \underline{\boldsymbol{\xi}}_k = \mathbf{M}\underline{\boldsymbol{\xi}}_{k-1} + \underline{\boldsymbol{\beta}}_k \\ \underline{\boldsymbol{\eta}}_k = \mathbf{N}\underline{\boldsymbol{\eta}}_{k-1} + \underline{\boldsymbol{\gamma}}_k \end{cases}$$

其中 \mathbf{M} 和 \mathbf{N} 是常值矩阵, $\{\underline{\boldsymbol{\beta}}_k\}$ 和 $\{\underline{\boldsymbol{\gamma}}_k\}$ 是独立的高斯白噪声序列, 且满足

$$E(\underline{\boldsymbol{\beta}}_k \underline{\boldsymbol{\beta}}_l^T) = \mathbf{Q}\delta_{kl}, \quad E(\underline{\boldsymbol{\gamma}}_k \underline{\boldsymbol{\gamma}}_l^T) = \mathbf{R}\delta_{kl}, \quad E(\underline{\boldsymbol{\beta}}_k \underline{\boldsymbol{\gamma}}_l^T) = \mathbf{0}$$

$$E(\mathbf{x}_0 \underline{\boldsymbol{\beta}}_l^T) = \mathbf{0}, \quad E(\mathbf{x}_0 \underline{\boldsymbol{\gamma}}_k^T) = \mathbf{0}, \quad \underline{\boldsymbol{\xi}}_{-1} = \underline{\boldsymbol{\eta}}_{-1} = \mathbf{0}$$

针对该系统的 α - β - γ - θ 跟踪器, 定义为下面的算法:

$$\begin{cases} \dot{\mathbf{X}}_k = \begin{bmatrix} \mathbf{A} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{N} \end{bmatrix} \dot{\mathbf{X}}_{k-1} + \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \theta \end{bmatrix} \left[\mathbf{v}_k - [\mathbf{C} \quad \mathbf{0} \quad \mathbf{I}] \begin{bmatrix} \mathbf{A} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{N} \end{bmatrix} \dot{\mathbf{X}}_{k-1} \right] \\ \dot{\mathbf{X}}_0 = \begin{bmatrix} E(\mathbf{x}_0) \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \end{cases}$$

式中 $\alpha, \beta, \gamma, \theta$ 是确定的常值, 且 $\mathbf{X}_k := [\mathbf{x}_k^T \quad \underline{\boldsymbol{\xi}}_k^T \quad \underline{\boldsymbol{\eta}}_k^T]^T$ 。

对于 9.2 节讨论的包含有色量测噪声输入的实时跟踪系统, 参数为

$$\mathbf{A} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{C} = [1 \quad 0 \quad 0]$$

$$\boldsymbol{\Gamma} = \mathbf{I}_3$$

$$\mathbf{Q} = \begin{bmatrix} \sigma_p & 0 & 0 \\ 0 & \sigma_v & 0 \\ 0 & 0 & \sigma_a \end{bmatrix}$$

$$R = [\sigma_m] > 0$$

$$M = 0$$

$$N = [r] > 0$$

其中 $h > 0$ 是采样时间, 当且仅当下面的条件满足时, 该 α - β - γ - θ 跟踪器是(近似最优)极限卡尔曼滤波: $\sigma_a > 0$, 且

- (1) $\gamma(2\alpha + 2\beta + \gamma) > 0$
 - (2) $\alpha(r-1)(r-1-r\theta) + r\beta\theta - r(r+1)\gamma\theta/2(r-1) > 0$
 - (3) $(4\alpha + \beta)(\beta + \gamma) + 3(\beta^2 - 2\alpha\gamma) - 4\gamma(r-1-\gamma\theta)/(r-1) \geq 0$
 - (4) $\gamma(\beta + \gamma) \geq 0$
 - (5) $\alpha(r+1)(r-1+r\theta) + r(r+1)\beta\theta/(r-1) + r\gamma(r+1)^2/2(r-1)^2 - r^2 + 1 \geq 0$
 - (6) $\sigma_v/\sigma_a = (\beta^2 - 2\alpha\gamma)/2h^2\gamma^2$
 - (7) $\sigma_p/\sigma_a = [\alpha(2\alpha + 2\beta + \gamma) - 4\beta(r-1-r\theta)/(r-1) - 4r\gamma\theta/(r-1)^2]h^4/(2\gamma^2)$
 - (8) $\sigma_m/\sigma_a = [\alpha(r+1)(r-1+r\theta) + 1 - r^2 + r^2\theta^2 + r(r+1)\beta\theta/(r-1) + r\gamma(r+1)^2/2(r-1)^2]h^4/\gamma^2$
 - (9) 矩阵 $[p_{ij}]_{4 \times 4}$ 是非负定对称的, 其中:
- $$p_{11} = \frac{1}{r-1} \left[\alpha(r-1-r\theta) + \frac{r\beta\theta}{r-1} - \frac{r\gamma\theta(r+1)}{2(r-1)^2} \right]$$
- $$p_{12} = \frac{1}{r-1} \left[\beta(r-1-r\theta) + \frac{r\gamma\theta}{r-1} \right]$$
- $$p_{13} = \frac{1}{r-1} \gamma(r-1-r\theta)$$
- $$p_{14} = \frac{r\theta}{r-1} \left[\alpha - \frac{\beta}{r-1} + \frac{\gamma(r+1)}{2(r-1)^2} \right]$$
- $$p_{22} = \frac{1}{4} (4\alpha + \beta)(\beta + \gamma) + \frac{3}{4} (\beta^2 - 2\alpha\gamma) - \frac{\gamma(r-1-r\theta)}{r-1}$$
- $$p_{23} = \gamma(\alpha + \beta/2)$$
- $$p_{24} = \frac{r\theta}{r-1} (\beta - \gamma/(r-1))$$
- $$p_{33} = \gamma(\beta + \gamma)$$
- $$p_{34} = r\gamma\theta/(r-1)$$
- $$p_{44} = \frac{1}{1-r^2} \left[\alpha(r+1)(r+r\theta-1) + \frac{r\beta\theta(r+1)}{r-1} + \frac{r\gamma(r+1)^2}{2(r-1)^2} + 1 - r^2 \right]$$

特别地, 如果 $N=0$, 也就是说, 只考虑白噪声输入, 则上面的条件简化为定理 9.1 描述的情况。

最后,通过定义 $\mathbf{x}_k := [x_k \quad \dot{x}_k \quad \ddot{x}_k \quad w_k]^T$, 并应用 z -变换, α - β - γ - θ 跟踪器可以分解为下面的(解耦)递推方程:

- (1) $x_k = a_1 x_{k-1} + a_2 x_{k-2} + a_3 x_{k-3} + a_4 x_{k-4} + \alpha v_k + (-2\alpha - r\alpha + \beta + \gamma/2)v_{k-1} + (\alpha - \beta + \gamma/2) + r(2\alpha - \beta - \gamma/2)v_{k-2} - r(\alpha - \beta + \gamma/2)v_{k-3},$
- (2) $\dot{x}_k = a_1 \dot{x}_{k-1} + a_2 \dot{x}_{k-2} + a_3 \dot{x}_{k-3} + a_4 \dot{x}_{k-4} + (1/h)\{\beta v_k - [(2+r)\beta - \gamma]v_{k-1} + [\beta - \gamma + r(2\beta - \gamma)]v_{k-2} - r(\beta - \gamma)v_{k-3}\}$
- (3) $\ddot{x}_k = a_1 \ddot{x}_{k-1} + a_2 \ddot{x}_{k-2} + a_3 \ddot{x}_{k-3} + \ddot{x}_{k-4} + (\gamma/h^2)[v_k - (2+\gamma)v_{k-1} + (1+2r)v_{k-2} - rv_{k-3}]$
- (4) $w_k = a_1 w_{k-1} + a_2 w_{k-2} + a_3 w_{k-3} + a_4 w_{k-4} + \theta(v_k - 3v_{k-1} + 3v_{k-2} - v_{k-3})$

初始条件为 x_{-1} 、 \dot{x}_{-1} 、 \ddot{x}_{-1} 和 w_{-1} , 其中:

$$\begin{cases} a_1 = -\alpha - \beta - \gamma/2 + r(\theta - 1) + 3 \\ a_2 = 2\alpha + \beta - \gamma/2 + r(\alpha + \beta + \gamma/2 + 3\theta - 3) - 3 \\ a_3 = -\alpha + r(-2\alpha - \beta + \gamma/2 - 3\theta + 3) + 1 \\ a_4 = r(\alpha + \theta - 1) \end{cases}$$

注意上面的解耦公式包含了9.2节得到的白噪声输入结果。此方法可以工作得很好,尽管滤波器只具有近似最优性。更详细的介绍见Chen和Chui(1986)和Chui(1984)。

12.3 自适应卡尔曼滤波

考虑线性随机状态空间模型:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \boldsymbol{\Gamma}_k \underline{\boldsymbol{\xi}}_k \\ \mathbf{v}_k = \mathbf{C}_k \mathbf{x}_k + \underline{\boldsymbol{\eta}}_k \end{cases}$$

式中 $\{\underline{\boldsymbol{\xi}}_k\}$ 和 $\{\underline{\boldsymbol{\eta}}_k\}$ 是不相关的零均值高斯白噪声序列,且有 $\text{Var}(\underline{\boldsymbol{\xi}}_k) = \mathbf{Q}_k$ 和 $\text{Var}(\underline{\boldsymbol{\eta}}_k) = \mathbf{R}_k$ 。然后假设每一个 \mathbf{R}_k 是正定的,所有的矩阵 \mathbf{A}_k 、 \mathbf{C}_k 、 $\boldsymbol{\Gamma}_k$ 、 \mathbf{Q}_k 和 \mathbf{R}_k ,以及初始条件 $E(\mathbf{x}_0)$ 和 $\text{Var}(\mathbf{x}_0)$ 是已知的。第2章和第3章推导的卡尔曼滤波算法提供了一个实时估计状态向量 \mathbf{x}_k 的非常有效的最优估计方案。事实上,该估计过程是如此的有效,以致选取的初始条件 $E(\mathbf{x}_0)$ 和/或 $\text{Var}(\mathbf{x}_0)$ 非常差时,也能够在较短的时间内获得较理想的估计值。然而,如果不是所有的矩阵 \mathbf{A}_k 、 \mathbf{C}_k 、 $\boldsymbol{\Gamma}_k$ 、 \mathbf{Q}_k 和 \mathbf{R}_k 都是已知值,滤波算法必须根据获得的数据 $\{\mathbf{v}_k\}$ 实时改变自己,使得状态的最优估计依然能够像这些矩阵都已知时一样获得。这类算法称为自

适应算法, 相应的滤波过程称为自适应卡尔曼滤波。如果 \mathbf{Q}_k 和 \mathbf{R}_k 是已知的, 则自适应卡尔曼滤波器可以用来辨识系统矩阵和/或量测矩阵。当 \mathbf{A}_k 、 \mathbf{C}_k 、 $\boldsymbol{\Gamma}_k$ 的部分信息已知时, 这个问题在第 8 章已经讨论过了。通常辨识问题是非常困难但又非常重要的[见 Aström 和 Eykhoff (1971) 及 Mehra(1970, 1972)]。

现在来讨论 \mathbf{A}_k 、 \mathbf{C}_k 、 $\boldsymbol{\Gamma}_k$ 已经给出的情况。与估计状态一样估计噪声方差矩阵的自适应卡尔曼滤波算法, 称为噪声自适应滤波器。虽然针对该问题有较多的算法[见 Aström 和 Eykhoff (1971), Chen 和 Chui (1991), Chin (1979), Jazwinski(1969) 和 Mehra(1970, 1972)], 至今还没有从最优准则的角度推导的算法。为了简便, 讨论 \mathbf{A}_k 、 \mathbf{C}_k 、 $\boldsymbol{\Gamma}_k$ 和 \mathbf{Q}_k 都给定的情形, 因此只有 \mathbf{R}_k 需要估计。新息过程[见 Kailath(1970) 和 Mehra(1970)]对此情形非常有效。从新获取的数据信息 v_k 和前一步获得的最优预测 $\hat{x}_{k|k-1}$, 新息序列定义为

$$\mathbf{z}_k := \mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}$$

显然

$$\mathbf{z}_k := \mathbf{C}_k (\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_{k|k-1}) + \underline{\eta}_k$$

实际上这是一个零均值白噪声序列。

对两边都取方差, 有

$$\mathbf{S}_k := \text{Var}(\mathbf{z}_k) = \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k$$

这就可以得到 \mathbf{R}_k 的估计, 即

$$\hat{\mathbf{R}}_k = \hat{\mathbf{S}}_k - \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T$$

其中 $\hat{\mathbf{S}}_k$ 是 \mathbf{S}_k 的统计采样方差

$$\hat{\mathbf{S}}_k = \frac{1}{k-1} \sum_{i=1}^k (\mathbf{z}_i - \bar{\mathbf{z}}_i) (\mathbf{z}_i - \bar{\mathbf{z}}_i)^T$$

其中 $\bar{\mathbf{z}}_i$ 是统计采样的均值

$$\bar{\mathbf{z}}_i = \frac{1}{i} \sum_{j=1}^i \mathbf{z}_j$$

(例如, 可参见 Stengel(1986))。

12.4 自适应卡尔曼滤波在维纳滤波中的应用

在数字信号处理和系统理论中, 一个重要的问题是根据受噪声污染的输入/输出信息确定数字滤波器或线性系统的单位脉冲响应, 具体例子见 Chui 和 Chen(1992)。更精确地, 设 $\{u_k\}$ 、 $\{v_k\}$ 分别是已知的输入、输出信号, $\{\eta_k\}$ 是未知的噪声序列。问题是下面的关系中“识别”出系数序列 $\{h_k\}$:

$$v_k = \sum_{i=0}^{\infty} h_i u_{k-i} + \eta_k, \quad k = 0, 1, \dots$$

在维纳滤波器中, 最优准则是确定序列 $\{\hat{h}_k\}$ 。通过设

$$\hat{v}_k = \sum_{i=0}^{\infty} \hat{h}_i u_{k-i}$$

需要满足

$$\text{Var}(v_k - \hat{v}_k) = \inf_{a_i} \left\{ \text{Var}(v_k - w_k) : \quad w_k = \sum_{i=0}^{\infty} a_i u_{k-i} \right\}.$$

在 $h_i = 0 (i > M)$ 的假设下, 也就是, 当考虑一个 FIR 系统, 上面的问题可以改写为下面的状态空间描述形式。设

$$\mathbf{x} = [h_0 \quad h_1 \quad \cdots \quad h_M]^T$$

是需要估计的状态向量。由于这是一个常值向量, 可以写为

$$\mathbf{x}_{k+1} = \mathbf{x}_k = \mathbf{x}$$

另设 \mathbf{C} 是“观测矩阵”, 定义为

$$\mathbf{C} = [u_0 \quad u_1 \quad \cdots \quad u_M]$$

然后输入/输出关系写为

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{x}_k \\ \mathbf{v}_k = \mathbf{C}\mathbf{x}_k + \eta_k \end{cases} \quad (\text{a})$$

需要从数据信息 $\{v_0, \dots, v_k\}$ 中给出 \mathbf{x}_k 的最优估计 $\hat{\mathbf{x}}_k$ 。当 $\{\eta_k\}$ 是具有未知方差 $R_k = \text{Var}(\eta_k)$ 的零均值高斯白噪声序列时, 该估计可以应用 12.3 节讨论的噪声自适应卡尔曼滤波完成。

对于 IIR 系统, 由于相应的线性系统(a)为无限维, 自适应卡尔曼滤波技术不能直接应用。

12.5 卡尔曼-布希滤波

这本书专门研究离散时间模型的卡尔曼滤波问题。下面将简单介绍连续时间模型问题, 称为卡尔曼-布希滤波器。

考虑连续时间线性确定/随机系统:

$$\begin{cases} d\mathbf{x}(t) = \mathbf{A}(t)\mathbf{x}(t)dt + \mathbf{B}(t)\mathbf{u}(t)dt + \boldsymbol{\Gamma}(t)\underline{\xi}(t)dt, \quad \mathbf{x}(0) = \mathbf{x}_0 \\ d\mathbf{v}(t) = \mathbf{C}(t)\mathbf{x}(t)dt + \underline{\eta}(t)dt \end{cases}$$

式中 $0 \leq t \leq T$ 。

状态向量 $\mathbf{x}(t)$ 是初值满足 $\mathbf{x}(0) \sim N(\mathbf{0}, \Sigma^2)$ 的随机 n 维向量, 这里 Σ^2 或者至少它的一个估计是给定的; 随机输入或噪声过程 $\underline{\xi}(t)$ 和 $\underline{\eta}(t)$ 分别是 p 维和 q 维 Weiner-Levy 向量, $1 \leq p, q \leq n$; 观测 $\mathbf{v}(t)$ 是 q 维随机向量; $\mathbf{A}(t), \Gamma(t)$ 和 $\mathbf{C}(t)$ 是连续时间区间 $[0, T]$ 上 $n \times n, n \times p$ 和 $q \times n$ 阶确定性连续矩阵函数。

卡尔曼-布希滤波由下面的递推公式给出

$$\begin{aligned}\hat{\mathbf{x}}(t) = & \int_0^T \mathbf{P}(\tau) \mathbf{C}^T(\tau) \mathbf{R}^{-1}(\tau) d\mathbf{v}(\tau) + \int_0^T [\mathbf{A}(\tau) - \mathbf{P}(\tau) \mathbf{C}^T(\tau) \mathbf{R}^{-1}(\tau) \mathbf{C}(\tau)] \hat{\mathbf{x}}(\tau) d(\tau) \\ & + \int_0^T \mathbf{B}(\tau) \mathbf{u}(\tau) d\tau\end{aligned}$$

式中 $\mathbf{R}(t) = E\{\underline{\eta}(t) \underline{\eta}^T(t)\}$, $\mathbf{P}(t)$ 满足 Riccati 方程

$$\begin{cases} \dot{\mathbf{P}}(t) = \mathbf{A}(t) \mathbf{P}(t) + \mathbf{P}(t) \mathbf{A}^T(t) - \mathbf{P}(t) \mathbf{C}^T(t) \mathbf{R}^{-1}(t) \mathbf{C}(t) \mathbf{P}(t) + \Gamma(t) \mathbf{Q}(t) \Gamma^T(t) \\ \mathbf{P}(0) = \text{Var}(\mathbf{x}_0) = \Sigma^2 \end{cases}$$

其中 $\mathbf{Q}(t) = E\{\underline{\xi}(t) \underline{\xi}^T(t)\}$ 。

想了解更多的细节, 读者可以阅读原始文献 Kalman 和 Bucy(1961), 教材 Ruymgaart 和 Soong(1995) 以及 Fleming 和 Rishel(1975)。

12.6 随机最优控制

关于确定性最优控制理论有大量的文献。关于该问题的简单讨论, 读者可以查阅文献 Chui 和 Chen(1989)。随机最优控制的目标是处理受随机扰动影响的系统。一个典型的随机最优控制是线性调节器问题, 它一直都受到广泛关注。我们将在本节讨论这一典型的问题。

系统和观测方程由线性确定性/随机微分方程给出:

$$\begin{cases} d\mathbf{x}(t) = \mathbf{A}(t) \mathbf{x}(t) dt + \mathbf{B}(t) \mathbf{u}(t) dt + \Gamma(t) \underline{\xi}(t) dt \\ d\mathbf{v}(t) = \mathbf{C}(t) \mathbf{x}(t) dt + \underline{\eta}(t) dt \end{cases}$$

式中 $0 \leq t \leq T$ (见 12.5 节), 定义在一类控制函数 $\mathbf{u}(t)$ 上的将要最小化的损失函数为

$$F(\mathbf{u}) = E \left\{ \int_0^T [\mathbf{x}^T(t) \mathbf{W}_x(t) \mathbf{x}(t) + \mathbf{u}^T(t) \mathbf{W}_u(t) \mathbf{u}(t)] dt \right\}$$

这里假设初始状态 $\mathbf{x}(0)$ 满足 $\mathbf{x}_0 \sim N(\mathbf{0}, \Sigma^2)$, Σ^2 是给定的, $\underline{\xi}(t)$ 和 $\underline{\eta}(t)$ 是不相关的高斯白噪声过程, 并且独立于初始状态 \mathbf{x}_0 。此外, 对于 $0 \leq t \leq T$, 数据

项 $v(t)$ 是已知的。 $A(t)$ 、 $B(t)$ 、 $C(t)$ 、 $W_x(t)$ 和 $W_u(t)$ 是已知的具有合适维数的确定矩阵， $W_x(t)$ 是非负定对称矩阵， $W_u(t)$ 是正定对称矩阵。通常，符合作为控制函数的 $u(t)$ 由定义在 $[0, T]$ 上的向量值 Borel 可观测函数组成，范围是 \mathbf{R}^p 上的一个闭子集。

假设控制函数 $u(t)$ 借助观测数据取得了系统状态的部分信息，也就是 $u(t)$ 是数据 $v(t)$ 而不是状态 $x(t)$ 的一个线性函数。对于这样的线性调节问题，可以应用所谓的分离原理，这是随机最优控制理论中最有用的结果之一。该原理本质上是指上面的“部分观测”线性调节问题可以分解为两部分：第一部分是应用 12.5 节讨论的卡尔曼-布希滤波器解决一个系统状态的最优估计问题；第二部分是一个“完全可观测”线性调节问题，该问题的解可以通过一个线性反馈函数给出。更精确地，状态向量 $x(t)$ 的最优估计 $\hat{x}(t)$ 满足线性随机系统：

$$\begin{cases} d\hat{x}(t) = A(t)\hat{x}(t)dt + B(t)u(t)dt + P(t)C^T(t)[dv(t) - R^{-1}(t)C(t)\hat{x}(t)dt] \\ \hat{x}(0) = E(x_0) \end{cases}$$

式中 $R(t) = E\{\underline{\eta}(t)\underline{\eta}^T(t)\}$, $P(t)$ 是下列矩阵 Riccati 方程的(唯一)解：

$$\begin{cases} \dot{P}(t) = A(t)P(t) + P(t)A^T(t) - P(t)C^T(t)R^{-1}(t)C(t)P(t) + \Gamma(t)Q(t)\Gamma^T(t) \\ P(0) = \text{Var}(x_0) = \Sigma \end{cases}$$

其中 $Q(t) = E\{\underline{\xi}(t)\underline{\xi}^T(t)\}$ 。

另一方面，最优控制函数 $u^*(t)$ 为

$$u^*(t) = -R^{-1}(t)B^T(t)K(t)\hat{x}(t)$$

式中 $K(t)$ 是下列矩阵 Riccati 方程的(唯一)解：

$$\begin{cases} \dot{K}(t) = K(t)B(t)W_u^{-1}(t)B^T(t)K(t) - K(t)A(t) - A^T(t)K(t) - W_x(t) \\ K(T) = 0 \end{cases}$$

其中 $0 \leq t \leq T$ 。

想了解更多的细节，读者可以阅读 Wonham(1968), Kushner(1971), Fleming 和 Rishel(1975), Davis(1977)，以及较新近的 Chen, Chen 和 Hsu(1995)。

12.7 平方根滤波及其脉动阵列实现

平方根滤波算法由 Potter(1963)首先提出，然后由 Carlson(1973)改进，它给出了一种快速计算方案。由第 7 章可以看出，该算法需要计算矩阵 $J_{k,k}$ 、

$\mathbf{J}_{k,k-1}$ 和 \mathbf{G}_k , 其中

$$\mathbf{J}_{k,k} = \mathbf{J}_{k,k-1} [\mathbf{I} - \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k + \mathbf{R}_k^c)^{-1} \mathbf{C}_k \mathbf{J}_{k,k-1}] \quad (a)$$

$\mathbf{J}_{k,k-1}$ 是下面矩阵的平方根:

$$[\mathbf{A}_{k-1} \mathbf{J}_{k-1,k-1} \quad \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}] [\mathbf{A}_{k-1} \mathbf{J}_{k-1,k-1} \quad \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}]^T$$

以及

$$\mathbf{G}_k = \mathbf{J}_{k,k-1} \mathbf{J}_{k,k-1}^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} \mathbf{H}_k^{-1}$$

其中, $\mathbf{J}_{k,k} = \mathbf{P}_{k,k-1}^{1/2}$ 、 $\mathbf{J}_{k,k-1} = \mathbf{P}_{k,k-1}^{1/2}$ 、 $\mathbf{H}_k = (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^c$, \mathbf{M}^c 是矩阵 \mathbf{M} 的下三角矩阵形式的“平方根”, 而不是 \mathbf{M} 是正定平方根 $\mathbf{M}^{1/2}$ (见引理 7.1)。很明显, 如果能够直接从 $\mathbf{J}_{k,k-1}$ 计算 $\mathbf{J}_{k,k}$ (或者直接从 $\mathbf{P}_{k,k-1}$ 计算 $\mathbf{P}_{k,k}$), 而不使用公式 (a), 该算法将更有效些。从这个角度出发, Bierman(1973, 1977)改进了 Carlson 的方法, 使用 LU 分解来给出下面的算法。

首先, 考虑下面的分解

$$\mathbf{P}_{k,k-1} = \mathbf{U}_1 \mathbf{D}_1 \mathbf{U}_1^T, \quad \mathbf{P}_{k,k} = \mathbf{U}_2 \mathbf{D}_2 \mathbf{U}_2^T$$

其中 \mathbf{U}_i 和 \mathbf{D}_i ($i=1, 2$) 分别是上三角矩阵和对角矩阵, 为了方便忽略了上标 k 。进一步, 定义

$$\mathbf{D} := \mathbf{D}_1 - \mathbf{D}_1 \mathbf{U}_1^T \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k)^{-1} \mathbf{C}_k \mathbf{U}_1 \mathbf{D}_1$$

并且分解 $\mathbf{D} = \mathbf{U}_3 \mathbf{D}_3 \mathbf{U}_3^T$, 则有

$$\mathbf{U}_2 = \mathbf{U}_1 \mathbf{U}_3, \quad \mathbf{D}_2 = \mathbf{D}_3$$

为了从 $\{\mathbf{U}_1, \mathbf{D}_1\}$ 得到 $\{\mathbf{U}_2, \mathbf{D}_2\}$, Bierman 的算法需要 $O(qn^2)$ 的算术操作, 其中 n 和 q 分别是状态向量和观测向量的维数。Andrews(1981)通过使用并行处理技术, 将操作减少为 $O(nq \cdot \log n)$ 。最近, Jover 和 Kailath(1986)使用 Schur 互补技术并应用脉动阵列(见 Kung(1982), Mead 和 Conway(1980) 和 Kung(1985))进一步将操作减少为 $O(n)$ (更精确地说, 近似为 $4n$)。此外, 需要的算术操作的数目从 $O(n^2)$ 减少为 $O(n)$ 。

该方案的基本思想可以简单地讨论如下: 因为 $\mathbf{P}_{k,k-1}$ 是非负定对称的, 存在正交矩阵 \mathbf{M}_1 , 使得

$$[\mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1}^{1/2} \quad \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1}^{1/2}] \mathbf{M}_1 = [\mathbf{P}_{k-1,k-1}^{1/2} \quad \mathbf{0}] \quad (b)$$

考虑增广矩阵

$$\mathbf{A} := \begin{bmatrix} \mathbf{H}_k \mathbf{H}_k^T & \mathbf{C}_k \mathbf{P}_{k,k-1} \\ \mathbf{P}_{k,k-1} \mathbf{C}_k^T & \mathbf{P}_{k,k-1} \end{bmatrix}$$

可以有以下两种分解:

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & \mathbf{C}_k \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{R}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{k,k} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C}_k^T & \mathbf{I} \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k)^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{H}_k \mathbf{H}_k^T & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{k,k} \end{bmatrix} \begin{bmatrix} \mathbf{I} & (\mathbf{H}_k^T)^{-1} (\mathbf{H}_k)^{-1} \mathbf{C}_k \mathbf{P}_{k,k-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

将分解的上三角平方根矩阵块放在左边, 分解的下三角平方根矩阵块放在右边, 另有一个正交矩阵 \mathbf{M}_2 , 使得

$$\begin{bmatrix} \mathbf{R}_k^{1/2} & \mathbf{C}_k \mathbf{P}_{k,k-1}^{1/2} \\ \mathbf{0} & \mathbf{P}_{k,k-1}^{1/2} \end{bmatrix} \mathbf{M}_2 = \begin{bmatrix} \mathbf{H}_k & \mathbf{0} \\ \mathbf{P}_{k,k-1} \mathbf{C}_k^T [(\mathbf{H}_k^T)^{-1/2} (\mathbf{H}_k)^{-1/2}]^T & \mathbf{P}_{k,k}^{1/2} \end{bmatrix} \quad (c)$$

使用 LU 分解, 为了简便, 省略上标 k , 有

$$\mathbf{R}_k = \mathbf{U}_R \mathbf{D}_R \mathbf{U}_R^T, \quad \mathbf{H}_k \mathbf{H}_k^T = \mathbf{U}_H \mathbf{D}_H \mathbf{U}_H^T$$

$$\mathbf{P}_{k,k-1} = \mathbf{U}_1 \mathbf{D}_1 \mathbf{U}_1^T, \quad \mathbf{P}_{k,k} = \mathbf{U}_2 \mathbf{D}_2 \mathbf{U}_2^T$$

恒等式(c)可以写为

$$\begin{bmatrix} \mathbf{U}_R & \mathbf{C}_k \mathbf{U}_1 \\ \mathbf{0} & \mathbf{U}_1 \end{bmatrix} \begin{bmatrix} \mathbf{D}_R & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_1 \end{bmatrix}^{1/2} \mathbf{M}_2 = \begin{bmatrix} \mathbf{U}_H & \mathbf{0} \\ \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{H}^T)^{-1} (\mathbf{H})^{-1} \mathbf{U}_H & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{D}_H & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix}^{1/2}$$

又定义

$$\mathbf{M}_3 = \begin{bmatrix} \mathbf{D}_R & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_1 \end{bmatrix}^{1/2} \mathbf{M}_2 \begin{bmatrix} \mathbf{D}_H & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix}^{-1/2}$$

很明显, 这是一个正交矩阵, 于是有

$$\begin{bmatrix} \mathbf{U}_R & \mathbf{C}_k \mathbf{U}_1 \\ \mathbf{0} & \mathbf{U}_1 \end{bmatrix} \mathbf{M}_3 = \begin{bmatrix} \mathbf{U}_H & \mathbf{0} \\ \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{H}^T)^{-1} (\mathbf{H})^{-1} \mathbf{U}_H & \mathbf{U}_2 \end{bmatrix} \quad (d)$$

应用 Kailath(1982)介绍的算法, \mathbf{M}_3 可以分解为有限个基本矩阵的乘积, 而不用平方根操作。所以, 通过合适的应用脉动阵列, 需要大约 $4n$ 的算术操作, $\{\mathbf{U}_H, \mathbf{U}_2\}$ 可以应用(d)式由 $\{\mathbf{U}_R, \mathbf{U}_1\}$ 得到, $\mathbf{P}_{k,k-1}^{1/2}$ 可以应用(b)由 $\mathbf{P}_{k-1,k-1}^{1/2}$ 得到。因此, \mathbf{D}_2 可以很容易从 \mathbf{D}_1 中计算得到。

要了解关于该主题的更多细节, 见 Jover 和 Kailath(1986) 及 Gaston 和 Irwin(1990)。

参 考 文 献

- Alfeld, G. and Herzberger, J. (1983): *Introduction to Interval Computations* (Academic, New York)
- Anderson, B.D.O., Moore, J.B. (1979): *Optimal Filtering* (Prentice-Hall, Englewood Cliffs, NJ)
- Andrews, A. (1981): "Parallel processing of the Kalman filter", IEEE Proc. Int. Conf. on Paral. Process., pp.216-220
- Aoki, M. (1989): *Optimization of Stochastic Systems: Topics in Discrete-Time Dynamics* (Academic, New York)
- Aström, K.J., Eykhoff, P. (1971): "System identification – a survey," *Automatica*, **7**, pp.123-162
- Balakrishnan, A.V. (1984,87): *Kalman Filtering Theory* (Optimization Software, Inc., New York)
- Bierman, G.J. (1973): "A comparison of discrete linear filtering algorithms," IEEE Trans. Aero. Elec. Systems, **9**, pp.28-37
- Bierman, G.J. (1977): *Factorization Methods for Discrete Sequential Estimation* (Academic, New York)
- Blahut, R.E. (1985): *Fast Algorithms for Digital Signal Processing* (Addison-Wesley, Reading, MA)
- Bozic, S.M. (1979): *Digital and Kalman Filtering* (Wiley, New York)
- Brammer, K., Sifflin, G. (1989): *Kalman-Bucy Filters* (Artech House, Boston)
- Brown, R.G. and Hwang, P.Y.C. (1992,97): *Introduction to Random Signals and Applied Kalman Filtering* (Wiley, New York)
- Bucy, R.S., Joseph, P.D. (1968): *Filtering for Stochastic Processes with Applications to Guidance* (Wiley, New York)
- Burrus, C.S. , Gopinath, R.A. and Guo, H. (1998): *Introduction to Wavelets and Wavelet Transfoms: A Primer* (Prentice-Hall, Upper Saddle River, NJ)
- Carlson, N.A. (1973): "Fast triangular formulation of the square root filter," J. ALAA, **11** pp.1259-1263
- Catlin, D.E. (1989): *Estimation, Control, and the Discrete Kalman Filter* (Springer, New York)
- Chen, G. (1992): "Convergence analysis for inexact mechanization of Kalman filtering," IEEE Trans. Aero. Elect. Syst., **28**,

pp.612-621

- Chen, G. (1993): *Approximate Kalman Filtering* (World Scientific, Singapore)
- Chen, G., Chen, G. and Hsu, S.H. (1995): *Linear Stochastic Control Systems* (CRC, Boca Raton, FL)
- Chen, G., Chui, C.K. (1986): "Design of near-optimal linear digital tracking filters with colored input," *J. Comp. Appl. Math.*, **15**, pp.353-370
- Chen, G., Wang, J. and Shieh, L.S. (1997): "Interval Kalman filtering," *IEEE Trans. Aero. Elect. Syst.*, **33**, pp.250-259
- Chen, H.F. (1985): *Recursive Estimation and Control for Stochastic Systems* (Wiley, New York)
- Chui, C.K. (1984): "Design and analysis of linear prediction-correction digital filters," *Linear and Multilinear Algebra*, **15**, pp.47-69
- Chui, C.K. (1997): *Wavelets: A Mathematical Tool for Signal Analysis*, (SIAM, Philadelphia)
- Chui, C.K., Chen, G. (1989): *Linear Systems and Optimal Control*, Springer Ser. Inf. Sci., Vol. 18 (Springer, Berlin Heidelberg)
- Chui, C.K., Chen, G. (1992, 1997): *Signal Processing and Systems Theory: Selected Topics*, Springer Ser. Inf. Sci., Vol. 26 (Springer, Berlin Heidelberg)
- Chui, C.K., Chen, G. and Chui, H.C. (1990): "Modified extended Kalman filtering and a real-time parallel algorithm for system parameter identification," *IEEE Trans. Auto. Control*, **35**, pp.100-104
- Davis, M.H.A. (1977): *Linear Estimation and Stochastic Control* (Wiley, New York)
- Davis, M.H.A., Vinter, R.B. (1985): *Stochastic Modeling and Control* (Chapman and Hall, New York)
- Fleming, W.H., Rishel, R.W. (1975): *Deterministic and Stochastic Optimal Control* (Springer, New York)
- Gaston, F.M.F., Irwin, G.W. (1990): "Systolic Kalman filtering: An overview," *IEE Proc.-D*, **137**, pp.235-244
- Goodwin, G.C., Sin, K.S. (1984): *Adaptive Filtering Prediction and Control* (Prentice-Hall, Englewood Cliffs, NJ)
- Haykin, S. (1986): *Adaptive Filter Theory* (Prentice-Hall, Englewood Cliffs, NJ)
- Hong, L., Chen, G. and Chui, C.K. (1998): "A filter-bank-based Kalman filtering technique for wavelet estimation and decom-

- position of random signals," IEEE Trans. Circ. Syst. (II), **45**, pp. 237-241.
- Hong, L., Chen, G. and Chui, C.K. (1998): "Real-time simultaneous estimation and decomposition of random signals," Multidim. Sys. Sign. Proc., **9**, pp. 273-289.
- Jazwinski, A.H. (1969): "Adaptive filtering," Automatica, **5**, pp.475-485
- Jazwinski, A.H. (1970): *Stochastic Processes and Filtering Theory* (Academic, New York)
- Jover, J.M., Kailath, T. (1986): "A parallel architecture for Kalman filter measurement update and parameter estimation," Automatica, **22**, pp.43-57
- Kailath, T. (1968): "An innovations approach to least-squares estimation, part I: linear filtering in additive white noise," IEEE Trans. Auto. Contr., **13**, pp.646-655
- Kailath, T. (1982): *Course Notes on Linear Estimation* (Stanford University, CA)
- Kalman, R.E. (1960): "A new approach to linear filtering and prediction problems," Trans. ASME, J. Basic Eng., **82**, pp.35-45
- Kalman, R.E. (1963): "New method in Wiener filtering theory," Proc. Symp. Eng. Appl. Random Function Theory and Probability (Wiley, New York)
- Kalman, R.E., Bucy, R.S. (1961): "New results in linear filtering and prediction theory," Trans. ASME J. Basic Eng., **83**, pp.95-108
- Kumar, P.R., Varaiya, P. (1986): *Stochastic Systems: Estimation, Identification, and Adaptive Control* (Prentice-Hall, Englewood Cliffs, NJ)
- Kung, H.T. (1982): "Why systolic architectures?" Computer, **15**, pp.37-46
- Kung, S.Y. (1985): "VLSI arrays processors," IEEE ASSP Magazine, **2**, pp.4-22
- Kushner, H. (1971): *Introduction to Stochastic Control* (Holt, Rinehart and Winston, Inc., New York)
- Lewis, F.L. (1986): *Optimal Estimation* (Wiley, New York)
- Lu, M., Qiao, X., Chen, G. (1992): "A parallel square-root algorithm for the modified extended Kalman filter," IEEE Trans. Aero. Elect. Syst., **28**, pp.153-163
- Lu, M., Qiao, X., Chen, G. (1993): "Parallel computation of the

- modified extended Kalman filter," Int'l J. Comput. Math., **45**, pp.69-87
- Maybeck, P.S. (1982): *Stochastic Models, Estimation, and Control*, Vol. 1,2,3 (Academic, New York)
- Mead, C., Conway, L. (1980): *Introduction to VLSI systems* (Addison-Wesley, Reading, MA)
- Mehra, R.K. (1970): "On the identification of variances and adaptive Kalman filtering," IEEE Trans. Auto. Contr., **15**, pp.175-184
- Mehra, R.K. (1972): "Approaches to adaptive filtering," IEEE Trans. Auto. Contr., **17**, pp.693-698
- Mendel, J.M. (1987): *Lessons in Digital Estimation Theory* (Prentice-Hall, Englewood Cliffs, New Jersey)
- Potter, J.E. (1963): "New statistical formulas," Instrumentation Lab., MIT, Space Guidance Analysis Memo. # 40
- Probability Group (1975), Institute of Mathematics, Academia Sinica, China (ed.): *Mathematical Methods of Filtering for Discrete-Time Systems* (in Chinese) (Beijing)
- Ruymgaart, P.A., Soong, T.T. (1985,88): *Mathematics of Kalman-Bucy Filtering*, Springer Ser. Inf. Sci., Vol. 14 (Springer, Berlin Heidelberg)
- Shiryayev, A.N. (1984): *Probability* (Springer-Verlag, New York)
- Siouris, G., Chen, G. and Wang, J. (1997): "Tracking an incoming ballistic missile," IEEE Trans. Aero. Elect. Syst., **33**, pp.232-240
- Sorenson, H.W., ed. (1985): *Kalman Filtering: Theory and Application* (IEEE, New York)
- Stengel, R.F. (1986): *Stochastic Optimal Control: Theory and Application* (Wiley, New York)
- Strobach, P. (1990): *Linear Prediction Theory: A Mathematical Basis for Adaptive Systems*, Springer Ser. Inf. Sci., Vol. 21 (Springer, Berlin Heidelberg)
- Wang, E.P. (1972): "Optimal linear recursive filtering methods," J. Mathematics in Practice and Theory (in Chinese), **6**, pp.40-50
- Wonham, W.M. (1968): "On the separation theorem of stochastic control," SIAM J. Control, **6**, pp.312-326
- Xu, J.H., Bian, G.R., Ni, C.K., Tang, G.X. (1981): *State Estimation and System Identification* (in Chinese) (Beijing)
- Young, P. (1984): *Recursive Estimation and Time-Series Analysis* (Springer, New York)

练习答案和提示

第 1 章

1.1 因为很多特性都可以直接由迹的定义证明, 只讨论 $\operatorname{tr}AB = \operatorname{tr}BA$:

$$\operatorname{tr}AB = \sum_{i=1}^n \left(\sum_{j=1}^m a_{ij} b_{ji} \right) = \sum_{j=1}^m \left(\sum_{i=1}^n b_{ji} a_{ij} \right) = \operatorname{tr}BA$$

$$1.2 (\operatorname{tr}A)^2 = \left(\sum_{i=1}^n a_{ii} \right)^2 \leq n \sum_{i=1}^n a_{ii}^2 \leq n(\operatorname{tr}AA^T)$$

$$1.3 A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

1.4 存在酉矩阵 P 和 Q , 使得

$$A = P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^T, \quad B = Q \begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{bmatrix} Q^T$$

及

$$\sum_{k=1}^n \lambda_k^2 \geq \sum_{k=1}^n \mu_k^2$$

设 $P = [p_{ij}]_{n \times n}$ 和 $Q = [q_{ij}]_{n \times n}$, 则

$$p_{11}^2 + p_{21}^2 + \cdots + p_{n1}^2 = 1, \quad p_{12}^2 + p_{22}^2 + \cdots + p_{n2}^2 = 1, \dots$$

$$p_{1n}^2 + p_{2n}^2 + \cdots + p_{nn}^2 = 1, \quad q_{11}^2 + q_{21}^2 + \cdots + q_{n1}^2 = 1$$

$$q_{12}^2 + q_{22}^2 + \cdots + q_{n2}^2 = 1, \quad \dots, \quad q_{1n}^2 + q_{2n}^2 + \cdots + q_{nn}^2 = 1$$

以及

$$\begin{aligned} \operatorname{tr}AA^T &= \operatorname{tr} \left\{ P \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{bmatrix} P^T \right\} \\ &= \operatorname{tr} \begin{bmatrix} p_{11}^2 \lambda_1^2 + p_{12}^2 \lambda_2^2 \\ * \\ * \\ p_{21}^2 \lambda_1^2 + p_{22}^2 \lambda_2^2 \\ * \\ * \\ p_{n1}^2 \lambda_1^2 + p_{n2}^2 \lambda_2^2 \\ * \\ * \end{bmatrix} \\ &\quad + \cdots + p_{1n}^2 \lambda_n^2 \end{aligned}$$

$$\begin{aligned} &= (p_{11}^2 + p_{21}^2 + \cdots + p_{n1}^2)\lambda_1^2 + \cdots + (p_{1n}^2 + p_{2n}^2 + \cdots + p_{nn}^2)\lambda_n^2 \\ &= \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2 \end{aligned}$$

同理, $\text{tr}BB^T = \mu_1^2 + \mu_2^2 + \cdots + \mu_n^2$ 。

可得 $\text{tr}AA^T \geq \text{tr}BB^T$ 。

1.5 定义: $I = \int_{-\infty}^{\infty} e^{-y^2} dy$ 。换用极坐标, 有

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right) \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy \\ &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta = \pi \end{aligned}$$

1.6 定义: $I(x) = \int_{-\infty}^{\infty} e^{-xy^2} dy$ 。由练习 1.5, 有

$$I(x) = \frac{1}{\sqrt{x}} \int_{-\infty}^{\infty} e^{-(\sqrt{xy})^2} d(\sqrt{xy}) = \sqrt{\pi/x}.$$

因此,

$$\int_{-\infty}^{\infty} y^2 e^{-y^2} dy = -\frac{d}{dx} I(x) \Big|_{x=1} = -\frac{d}{dx} (\sqrt{\pi/x}) \Big|_{x=1} = \frac{1}{2} \sqrt{\pi}.$$

1.7 (a) 设 \mathbf{P} 是酉矩阵, 满足

$$\mathbf{R} = \mathbf{P}^T \text{diag}[\lambda_1, \dots, \lambda_n] \mathbf{P}$$

并定义: $\mathbf{y} = \frac{1}{\sqrt{2}} \text{diag}[\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}] \mathbf{P}(\mathbf{x} - \underline{\boldsymbol{\mu}})$ 。则

$$\begin{aligned} E(\mathbf{X}) &= \int_{-\infty}^{\infty} \mathbf{x} f(\mathbf{x}) dx \\ &= \int_{-\infty}^{\infty} (\underline{\boldsymbol{\mu}} + \sqrt{2} \mathbf{P}^{-1} \text{diag}[1/\sqrt{\lambda_1}, \dots, 1/\sqrt{\lambda_n}] \mathbf{y}) f(\mathbf{x}) dx \\ &= \underline{\boldsymbol{\mu}} \int_{-\infty}^{\infty} f(\mathbf{x}) dx + \text{Const.} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} e^{-y_1^2} \cdots e^{-y_n^2} dy_1 \cdots dy_n \\ &= \underline{\boldsymbol{\mu}} \cdot 1 + \mathbf{0} = \underline{\boldsymbol{\mu}} \end{aligned}$$

(b) 使用同样的代换, 有

$$\begin{aligned} \text{Var}(\mathbf{X}) &= \int_{-\infty}^{\infty} (\mathbf{x} - \underline{\boldsymbol{\mu}})(\mathbf{x} - \underline{\boldsymbol{\mu}})^T f(\mathbf{x}) dx = \int_{-\infty}^{\infty} 2\mathbf{R}^{1/2} \mathbf{y} \mathbf{y}^T \mathbf{R}^{1/2} f(\mathbf{x}) dx \\ &= \frac{2}{(\pi)^{n/2}} \mathbf{R}^{1/2} \left\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \begin{bmatrix} y_1^2 & \cdots & y_1 y_n \\ \vdots & & \vdots \\ y_n y_1 & \cdots & y_n^2 \end{bmatrix} e^{-y_1^2} \cdots e^{-y_n^2} dy_1 \cdots dy_n \right\} \mathbf{R}^{1/2} \\ &= \mathbf{R}^{1/2} \mathbf{I} \mathbf{R}^{1/2} = \mathbf{R} \end{aligned}$$

1.8 所有的性质可以很容易地从定义得到。

1.9 已证明如果 X_1, X_2 相互独立, 则 $\text{cov}(X_1, X_2) = \mathbf{0}$ 。假设 $\mathbf{R}_{12} = \text{cov}(X_1, X_2) = \mathbf{0}$, 则 $\mathbf{R}_{21} = \text{cov}(X_2, X_1) = \mathbf{0}$, 使得

$$\begin{aligned} f(X_1, X_2) &= \frac{1}{(2\pi)^{n/2} \det \mathbf{R}_{11} \det \mathbf{R}_{22}} e^{-\frac{1}{2}(X_1 - \underline{\mu}_1)^T \mathbf{R}_{11} (X_1 - \underline{\mu}_1)} e^{-\frac{1}{2}(X_2 - \underline{\mu}_2)^T \mathbf{R}_{22} (X_2 - \underline{\mu}_2)} \\ &= f_1(X_1) f_2(X_2) \end{aligned}$$

所以, X_1 和 X_2 相互独立。

1.10 通过直接计算可以验证式(1.35)。首先, 有

$$\begin{bmatrix} \mathbf{I} & -\mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{R}_{yy}^{-1} \mathbf{R}_{xy}^T & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{xx} - \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{yy} \end{bmatrix}$$

然后, 通过取行列式得到

$$\begin{aligned} \det \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_{yy} \end{bmatrix} &= \det [\mathbf{R}_{xx} - \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx}] \det \mathbf{R}_{yy} \\ &= \left[\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} \underline{\mu}_x \\ \underline{\mu}_y \end{bmatrix} \right]^T \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_{yy} \end{bmatrix}^{-1} \left[\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} \underline{\mu}_x \\ \underline{\mu}_y \end{bmatrix} \right] \\ &= (\mathbf{x} - \bar{\underline{\mu}})^T [\mathbf{R}_{xx} - \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx}]^{-1} (\mathbf{x} - \bar{\underline{\mu}}) + (\mathbf{y} - \bar{\underline{\mu}}_y)^T \mathbf{R}_{yy}^{-1} (\mathbf{y} - \bar{\underline{\mu}}_y) \end{aligned}$$

式中

$$\bar{\underline{\mu}} = \underline{\mu}_x + \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} (\mathbf{y} - \bar{\underline{\mu}}_y)$$

剩下的计算步骤可直接得到。

1.11 设 $\mathbf{p}_k = \mathbf{C}_k^T \mathbf{W}_k \mathbf{z}_k$ 和 $\sigma^2 = E[\mathbf{p}_k^T (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{p}_k]$ 。可以证明:

$$F(\mathbf{y}_k) = \mathbf{y}_k^T (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k) \mathbf{y}_k - \mathbf{p}_k^T \mathbf{y}_k - \mathbf{y}_k^T \mathbf{p}_k + \sigma^2$$

又

$$\frac{dF(\mathbf{y}_k)}{d\mathbf{y}_k} = 2(\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k) \mathbf{y}_k - 2\mathbf{p}_k = \mathbf{0}$$

并假设矩阵 $(\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)$ 非奇异, 有

$$\hat{\mathbf{y}}_k = (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{p}_k = (\mathbf{C}_k^T \mathbf{W}_k \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{W}_k \mathbf{z}_k$$

$$\begin{aligned} 1.12 \quad E \hat{\mathbf{x}}_k &= (\mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{R}_k^{-1} E(\mathbf{v}_k - \mathbf{D}_k \mathbf{u}_k) \\ &= (\mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{R}_k^{-1} E(\mathbf{C}_k \mathbf{x}_k + \underline{\eta}_k) = E \mathbf{x}_k \end{aligned}$$

第 2 章

$$\begin{aligned} 2.1 \quad \mathbf{W}_{k,k-1}^{-1} &= \text{Var}(\bar{\underline{\epsilon}}_{k,k-1}) = E(\bar{\underline{\epsilon}}_{k,k-1} \bar{\underline{\epsilon}}_{k,k-1}^T) \\ &= E(\bar{\mathbf{v}}_{k-1} - \mathbf{H}_{k,k-1} \mathbf{x}_k)(\bar{\mathbf{v}}_{k-1} - \mathbf{H}_{k,k-1} \mathbf{x}_k)^T \end{aligned}$$

$$= \begin{bmatrix} \mathbf{R}_0 & & \\ & \ddots & \\ & & \mathbf{R}_{k-1} \end{bmatrix} + \text{Var} \begin{bmatrix} \mathbf{C}_0 \sum_{i=1}^k \boldsymbol{\Phi}_{0i} \boldsymbol{\Gamma}_{i-1} \underline{\boldsymbol{\xi}}_{i-1} \\ \vdots \\ \mathbf{C}_{k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1} \underline{\boldsymbol{\xi}}_{k-1} \end{bmatrix}$$

2.2 对于任意非零向量 \mathbf{x} , 有 $\mathbf{x}^\top \mathbf{A}\mathbf{x} > 0$ 和 $\mathbf{x}^\top \mathbf{B}\mathbf{x} \geq 0$, 使得 $\mathbf{x}^\top (\mathbf{A} + \mathbf{B})\mathbf{x} = \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{x}^\top \mathbf{B}\mathbf{x} > 0$, 所以 $\mathbf{A} + \mathbf{B}$ 是正定的。

$$\begin{aligned} 2.3 \quad & \mathbf{W}_{k,k-1}^{-1} \\ &= E(\underline{\boldsymbol{\epsilon}}_{k,k-1} \underline{\boldsymbol{\epsilon}}_{k,k-1}^\top) \\ &= E(\underline{\boldsymbol{\epsilon}}_{k-1,k-1} - \mathbf{H}_{k,k-1} \boldsymbol{\Gamma}_{k-1} \underline{\boldsymbol{\xi}}_{k-1})(\underline{\boldsymbol{\epsilon}}_{k-1,k-1} - \mathbf{H}_{k,k-1} \boldsymbol{\Gamma}_{k-1} \underline{\boldsymbol{\xi}}_{k-1})^\top \\ &= E(\underline{\boldsymbol{\epsilon}}_{k-1,k-1} \underline{\boldsymbol{\epsilon}}_{k-1,k-1}^\top) + \mathbf{H}_{k,k-1} \boldsymbol{\Gamma}_{k-1} E(\underline{\boldsymbol{\xi}}_{k-1} \underline{\boldsymbol{\xi}}_{k-1}^\top) \boldsymbol{\Gamma}_{k-1}^\top \mathbf{H}_{k,k-1}^\top \\ &= \mathbf{W}_{k-1,k-1}^{-1} + \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^\top \boldsymbol{\Phi}_{k-1,k}^\top \mathbf{H}_{k-1,k-1}^\top \end{aligned}$$

2.4 对 $\mathbf{A}_{11} = \mathbf{W}_{k-1,k-1}^{-1}, \mathbf{A}_{22} = \mathbf{Q}_{k-1}^{-1}$ 和 $\mathbf{A}_{12} = \mathbf{A}_{21}^\top = \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1}$ 应用引理 1.2。

2.5 由练习 2.4 或式(2.9), 有

$$\begin{aligned} & \mathbf{H}_{k,k-1}^\top \mathbf{W}_{k,k-1} \\ &= \boldsymbol{\Phi}_{k-1,k}^\top \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \\ &\quad - \boldsymbol{\Phi}_{k-1,k}^\top \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \mathbf{H}_{k,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1} \\ &\quad \cdot (\mathbf{Q}_{k-1}^{-1} + \boldsymbol{\Gamma}_{k-1}^\top \boldsymbol{\Phi}_{k-1,k}^\top \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1})^{-1} \\ &\quad \cdot \boldsymbol{\Gamma}_{k-1}^\top \boldsymbol{\Phi}_{k-1,k}^\top \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \\ &= \boldsymbol{\Phi}_{k-1,k}^\top \{ \mathbf{I} - \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1} \\ &\quad \cdot (\mathbf{Q}_{k-1}^{-1} + \boldsymbol{\Gamma}_{k-1}^\top \boldsymbol{\Phi}_{k-1,k}^\top \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1})^{-1} \\ &\quad \cdot \boldsymbol{\Gamma}_{k-1}^\top \boldsymbol{\Phi}_{k-1,k}^\top \} \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \end{aligned}$$

2.6 由练习 2.5 或式(2.10), 以及恒等式 $\mathbf{H}_{k,k-1} = \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k}$, 有

$$\begin{aligned} & (\mathbf{H}_{k,k-1}^\top \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1}) \boldsymbol{\Phi}_{k,k-1} \\ &\quad \cdot (\mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1})^{-1} \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \\ &= \boldsymbol{\Phi}_{k-1,k}^\top \{ \mathbf{I} - \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1} \\ &\quad \cdot (\mathbf{Q}_{k-1}^{-1} + \boldsymbol{\Gamma}_{k-1}^\top \boldsymbol{\Phi}_{k-1,k}^\top \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \boldsymbol{\Phi}_{k-1,k} \boldsymbol{\Gamma}_{k-1})^{-1} \\ &\quad \cdot \boldsymbol{\Gamma}_{k-1}^\top \boldsymbol{\Phi}_{k-1,k}^\top \} \mathbf{H}_{k-1,k-1}^\top \mathbf{W}_{k-1,k-1} \\ &= \mathbf{H}_{k,k-1}^\top \mathbf{W}_{k,k-1} \end{aligned}$$

$$\begin{aligned} 2.7 \quad & \mathbf{P}_{k,k-1} \mathbf{C}_k^\top (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^\top + \mathbf{R}_k)^{-1} \\ &= \mathbf{P}_{k,k-1} \mathbf{C}_k^\top (\mathbf{R}_k^{-1} - \mathbf{R}_k^{-1} \mathbf{C}_k (\mathbf{P}_{k,k-1}^{-1} + \mathbf{C}_k^\top \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1} \mathbf{C}_k^\top \mathbf{R}_k^{-1}) \\ &= (\mathbf{P}_{k,k-1} - \mathbf{P}_{k,k-1} \mathbf{C}_k^\top \mathbf{R}_k^{-1} \mathbf{C}_k (\mathbf{P}_{k,k-1}^{-1} + \mathbf{C}_k^\top \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1}) \mathbf{C}_k^\top \mathbf{R}_k^{-1} \\ &= (\mathbf{P}_{k,k-1} - \mathbf{P}_{k,k-1} \mathbf{C}_k^\top (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^\top + \mathbf{R}_k)^{-1}) \end{aligned}$$

$$\begin{aligned}
& \cdot (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k) \mathbf{R}_k^{-1} \mathbf{C}_k (\mathbf{P}_{k,k-1}^{-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{R}_k^{-1} \\
= & (\mathbf{P}_{k,k-1} - \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\
& \cdot (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k + \mathbf{C}_k) (\mathbf{P}_{k,k-1}^{-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1}) \mathbf{C}_k^T \mathbf{R}_k^{-1} \\
= & (\mathbf{P}_{k,k-1} - \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \mathbf{C}_k \mathbf{P}_{k,k-1} \\
& \cdot (\mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k + \mathbf{P}_{k,k-1}^{-1}) (\mathbf{P}_{k,k-1}^{-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k)^{-1}) \mathbf{C}_k^T \mathbf{R}_k^{-1} \\
= & (\mathbf{P}_{k,k-1} - \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \mathbf{C}_k \mathbf{P}_{k,k-1}) \mathbf{C}_k^T \mathbf{R}_k^{-1} \\
= & \mathbf{P}_{k,k} \mathbf{C}_k^T \mathbf{R}_k^{-1} \\
= & \mathbf{G}_k
\end{aligned}$$

2. 8 $\mathbf{P}_{k,k-1}$

$$\begin{aligned}
& = (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1} \\
& = (\Phi_{k-1,k}^T (\mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \\
& \quad - \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1}) \\
& \quad \cdot (\mathbf{Q}_{k-1}^{-1} + \Gamma_{k-1}^T \Phi_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\
& \quad \cdot \Gamma_{k-1}^T \Phi_{k-1,k}^T \mathbf{H}_{k-1,k-1}^T \mathbf{W}_{k-1,k-1} \mathbf{H}_{k-1,k-1} \Phi_{k-1,k})^{-1} \\
& = (\Phi_{k-1,k}^T \mathbf{P}_{k-1,k-1}^{-1} \Phi_{k-1,k} - \Phi_{k-1,k}^T \mathbf{P}_{k-1,k-1}^{-1} \Phi_{k-1,k} \Gamma_{k-1}) \\
& \quad \cdot (\mathbf{Q}_{k-1}^{-1} + \Gamma_{k-1}^T \Phi_{k-1,k}^T \mathbf{P}_{k-1,k-1}^{-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\
& \quad \cdot \Gamma_{k-1}^T \Phi_{k-1,k}^T \mathbf{P}_{k-1,k-1}^{-1} \Phi_{k-1,k})^{-1} \\
& = (\Phi_{k-1,k}^T \mathbf{P}_{k-1,k-1}^{-1} \Phi_{k-1,k})^{-1} + \Gamma_{k-1} \mathbf{Q}_{k-1} \Gamma_{k-1}^T \\
& = \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \Gamma_{k-1} \mathbf{Q}_{k-1} \Gamma_{k-1}^T
\end{aligned}$$

2. 9 $E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T$

$$\begin{aligned}
& = E(\mathbf{x}_k - (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1} \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \bar{\mathbf{v}}_{k-1}) \\
& \quad \cdot (\mathbf{x}_k - (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1} \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \bar{\mathbf{v}}_{k-1})^T \\
& = E(\mathbf{x}_k - (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1} \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \\
& \quad \cdot (\mathbf{H}_{k,k-1} \mathbf{x}_k + \underline{\boldsymbol{\epsilon}}_{k,k-1})) (\mathbf{x}_k - (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1} \\
& \quad \cdot \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} (\mathbf{H}_{k,k-1} \mathbf{x}_k + \underline{\boldsymbol{\epsilon}}_{k,k-1}))^T \\
& = (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1} \mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} E(\underline{\boldsymbol{\epsilon}}_{k,k-1} \underline{\boldsymbol{\epsilon}}_{k,k-1}^T) \mathbf{W}_{k,k-1} \\
& \quad \cdot \mathbf{H}_{k,k-1} (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1} \\
& = (\mathbf{H}_{k,k-1}^T \mathbf{W}_{k,k-1} \mathbf{H}_{k,k-1})^{-1} \\
& = \mathbf{P}_{k,k-1}
\end{aligned}$$

第二个等式的推导与上式类似。

2. 10 因为

$$\begin{aligned}
\sigma^2 & = \text{Var}(x_k) = E(ax_{k-1} + \xi_{k-1})^2 \\
& = a^2 \text{Var}(x_{k-1}) + 2aE(x_{k-1}\xi_{k-1}) + E(\xi_{k-1}^2) \\
& = a^2 \sigma^2 + \mu^2
\end{aligned}$$

有 $\sigma^2 = \mu^2 / (1 - a^2)$ 。

对 $j=1$, 有

$$E(x_k x_{k+1}) = E(x_k (ax_k + \xi_k)) = a \text{Var}(x_k) + E(x_k \xi_k) = a\sigma^2$$

对 $j=2$, 有

$$\begin{aligned} E(x_k x_{k+2}) &= E(x_k (ax_{k+1} + \xi_{k+1})) = aE(x_k x_{k+1}) + E(x_k + \xi_{k+1}) \\ &= aE(x_k x_{k+1}) = a^2\sigma^2 \end{aligned}$$

等。如果 j 为负数, 也可以得到类似的结果。

由归纳法, 可以得出, 对于所有的整数 j , 有 $E(x_k x_{k+j}) = a^{|j|}\sigma^2$ 。

2.11 由卡尔曼滤波方程(2.17), 有

$$P_{0,0} = \text{Var}(x_0) = \mu^2$$

$$P_{k,k-1} = P_{k-1,k-1}$$

$$G_k = P_{k,k-1} (P_{k,k-1} + R_k)^{-1} = \frac{P_{k-1,k-1}}{P_{k-1,k-1} + \sigma^2}$$

$$P_{k,k} = (1 - G_k) P_{k,k-1} = \frac{\sigma^2 P_{k-1,k-1}}{\sigma^2 + P_{k-1,k-1}}$$

又

$$P_{1,1} = \frac{\sigma^2 \mu^2}{\mu^2 + \sigma^2}$$

$$P_{2,2} = \frac{\sigma^2 P_{1,1}}{P_{1,1} + \sigma^2} = \frac{\sigma^2 \mu^2}{2\mu^2 + \sigma^2}$$

\vdots

$$P_{k,k} = \frac{\sigma^2 \mu^2}{k\mu^2 + \sigma^2}$$

因此

$$G_k = \frac{P_{k-1,k-1}}{P_{k-1,k-1} + \sigma^2} = \frac{\mu^2}{k\mu^2 + \sigma^2}$$

使得

$$\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + G_k (v_k - \hat{x}_{k|k-1}) \\ &= \hat{x}_{k-1|k-1} + \frac{\mu^2}{\sigma^2 + k\mu^2} (v_k - \hat{x}_{k-1|k-1}) \end{aligned}$$

初值为 $\hat{x}_{0|0} = E(x_0) = 0$ 。

对于值较大的 k , 有

$$\hat{x}_{k|k} = \hat{x}_{k-1|k-1}$$

$$2.12 \quad \hat{Q}_N = \frac{1}{N} \sum_{k=1}^N (\mathbf{v}_k \mathbf{v}_k^T) = \frac{1}{N} (\mathbf{v}_N \mathbf{v}_N^T) + \frac{1}{N} \sum_{k=1}^{N-1} (\mathbf{v}_k \mathbf{v}_k^T)$$

$$= \frac{1}{N}(\mathbf{v}_N \mathbf{v}_N^T) + \frac{N-1}{N} \hat{\mathbf{Q}}_{N-1} = \hat{\mathbf{Q}}_{N-1} + \frac{1}{N}[(\mathbf{v}_N \mathbf{v}_N^T) - \hat{\mathbf{Q}}_{N-1}]$$

初始估计为 $\hat{\mathbf{Q}}_1 = \mathbf{v}_1 \mathbf{v}_1^T$ 。

2.13 应用叠加原理。

2.14 设对于每一个 $k=0, 1, \dots, k$, $\mathbf{x}_k = [(\mathbf{x}_k^1)^T \dots (\mathbf{x}_k^N)^T]^T$, 对于所有的 $j < 0$, 有 $\mathbf{x}_j = \mathbf{0}$ (和 $\mathbf{u}_j = \mathbf{0}$), 并定义

$$\begin{aligned}\mathbf{x}_k^1 &= \mathbf{B}_1 \mathbf{x}_{k-1}^1 + \mathbf{x}_{k-1}^2 + (\mathbf{A}_1 + \mathbf{B}_1 \mathbf{A}_0) \mathbf{u}_{k-1} \\ &\vdots \\ \mathbf{x}_k^M &= \mathbf{B}_M \mathbf{x}_{k-1}^1 + \mathbf{x}_{k-1}^{M+1} + (\mathbf{A}_M + \mathbf{B}_M \mathbf{A}_0) \mathbf{u}_{k-1} \\ \mathbf{x}_k^{M+1} &= \mathbf{B}_{M+1} \mathbf{x}_{k-1}^1 + \mathbf{x}_{k-1}^{M+2} + \mathbf{B}_{M+1} \mathbf{A}_0 \mathbf{u}_{k-1} \\ &\vdots \\ \mathbf{x}_k^{N-1} &= \mathbf{B}_{N-1} \mathbf{x}_{k-1}^1 + \mathbf{x}_{k-1}^N + \mathbf{B}_{N-1} \mathbf{A}_0 \mathbf{u}_{k-1} \\ \mathbf{x}_k^N &= \mathbf{B}_N \mathbf{x}_{k-1}^1 + \mathbf{B}_N \mathbf{A}_0 \mathbf{u}_{k-1}\end{aligned}$$

将这些方程代入

$$\mathbf{v}_k = \mathbf{C} \mathbf{x}_k + \mathbf{D} \mathbf{u}_k = \mathbf{x}_k^1 + \mathbf{A}_0 \mathbf{u}_k$$

就可得到期望的结果。

因为对于所有 $j < 0$, 有 $\mathbf{x}_j = \mathbf{0}$ 和 $\mathbf{u}_j = \mathbf{0}$, 可得 $\mathbf{x}_0 = \mathbf{0}$ 。

第3章

3.1 设 $\mathbf{A} = \mathbf{B} \mathbf{B}^T$, $\mathbf{B} = [b_{ij}] \neq \mathbf{0}$ 。则 $\text{tr} \mathbf{A} = \text{tr} \mathbf{B} \mathbf{B}^T = \sum_{i,j} b_{ij}^2 > 0$ 。

3.2 根据假设 2.1, 由于 $l \geq j$, $\underline{\eta}_l$ 与 $\mathbf{x}_0, \underline{\xi}_0, \dots, \underline{\xi}_{j-1}, \underline{\eta}_0, \dots, \underline{\eta}_{j-1}$ 独立。

对于常值矩阵 $\mathbf{B}_0, \mathbf{B}_{1i}$ 和 \mathbf{B}_{2i} , 有

$$\begin{aligned}\hat{\mathbf{e}}_j &= \mathbf{C}_j (\mathbf{x}_j - \hat{\mathbf{y}}_{j-1}) \\ &= \mathbf{C}_j \left(\mathbf{A}_{j-1} \mathbf{x}_{j-1} + \mathbf{F}_{j-1} \underline{\xi}_{j-1} - \sum_{i=0}^{j-1} \hat{\mathbf{P}}_{j-1,i} (\mathbf{C}_i \mathbf{x}_i + \underline{\eta}_i) \right) \\ &\quad \cdots \\ &= \mathbf{B}_0 \mathbf{x}_0 + \sum_{i=0}^{j-1} \mathbf{B}_{1i} \underline{\xi}_i + \sum_{i=0}^{j-1} \mathbf{B}_{2i} \underline{\eta}_i\end{aligned}$$

从而对于所有的 $l \geq j$, 有 $\langle \underline{\eta}_l, \hat{\mathbf{e}}_j \rangle = \mathbf{O}_{q \times q}$ 。

3.3 联立式 (3.8) 和式 (3.4), 有 $\mathbf{e}_j = \|\mathbf{z}_j\|_q^{-1} \mathbf{z}_j = \|\mathbf{z}_j\|_q^{-1} \mathbf{v}_j - \sum_{i=0}^{j-1} (\|\mathbf{z}_j\|_q^{-1} \mathbf{C}_j \hat{\mathbf{P}}_{j-1,i}) \mathbf{v}_i$, 也就是说, \mathbf{e}_j 可以用 $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_j$ 表示。

反过来, 有

$$\mathbf{v}_0 = \mathbf{z}_0 = \|\mathbf{z}_0\|_q \mathbf{e}_0$$

$$\mathbf{v}_1 = \mathbf{z}_1 + \mathbf{C}_1 \hat{\mathbf{y}}_0 = \mathbf{z}_1 + \mathbf{C}_1 \hat{\mathbf{P}}_{0,0} \mathbf{v}_0 = \| \mathbf{z}_1 \|_q \mathbf{e}_1 + \mathbf{C}_1 \hat{\mathbf{P}}_{0,0} \| \mathbf{z}_0 \|_q \mathbf{e}_0$$

等等;也就是说 \mathbf{v}_j 也可以用 $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_j$ 来表示,所以有 $Y(\mathbf{e}_0, \dots, \mathbf{e}_k) = Y(\mathbf{v}_0, \dots, \mathbf{v}_k)$ 。

3.4 由练习3.3,对一些 $q \times q$ 阶常值矩阵 $\mathbf{L}_l, l=0, 1, \dots, i, i=0, 1, \dots, k-1$ 有

$$\mathbf{v}_i = \sum_{l=0}^i \mathbf{L}_l \mathbf{e}_l$$

使得

$$\langle \mathbf{v}_i, \mathbf{z}_k \rangle = \sum_{l=0}^i \mathbf{L}_l \langle \mathbf{e}_l, \mathbf{e}_k \rangle \| \mathbf{z}_k \|_q^T = \mathbf{O}_{q \times q}$$

则对于 $j=0, 1, \dots, k-1$,有 $\langle \hat{\mathbf{y}}_j, \mathbf{z}_k \rangle = \left\langle \sum_{i=0}^j \hat{\mathbf{P}}_{j,i} \mathbf{v}_i, \mathbf{z}_k \right\rangle = \sum_{i=0}^j \hat{\mathbf{P}}_{j,i} \langle \mathbf{v}_i, \mathbf{z}_k \rangle = \mathbf{O}_{n \times q}$ 。

3.5 因为

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}_{k-1} \mathbf{x}_{k-1} + \boldsymbol{\Gamma}_{k-1} \underline{\xi}_{k-1} \\ &= \mathbf{A}_{k-1} (\mathbf{A}_{k-2} \mathbf{x}_{k-2} + \boldsymbol{\Gamma}_{k-2} \underline{\xi}_{k-2}) + \boldsymbol{\Gamma}_{k-1} \underline{\xi}_{k-1} \\ &= \cdots = \mathbf{B}_0 \mathbf{x}_0 + \sum_{i=0}^{k-1} \mathbf{B}_{1i} \underline{\xi}_i \end{aligned}$$

其中 $\mathbf{B}_0, \mathbf{B}_{1i}$ 为常值矩阵,并且 $\underline{\xi}_k$ 独立于 \mathbf{x}_0 和 $\underline{\xi}_i (0 \leq i \leq k-1)$,于是有 $\langle \mathbf{x}_k, \underline{\xi}_k \rangle = \mathbf{0}$ 。剩下的可以同理得出。

3.6 应用叠加原理。

3.7 应用在练习3.6中得到的公式,有

$$\begin{cases} \hat{d}_{k|k} = \hat{d}_{k-1|k-1} + h w_{k-1} + G_k (\mathbf{v}_k - \Delta d_k - \hat{d}_{k-1|k-1} - h w_{k-1}) \\ \hat{d}_{0|0} = E(d_0) \end{cases}$$

式中 G_k 由标准算法(3.25)得出,且 $A_k = C_k = \boldsymbol{\Gamma}_k = 1$ 。

3.8 设

$$\begin{aligned} \mathbf{X}_k &= \begin{bmatrix} \mathbf{x}_k^1 \\ \mathbf{x}_k^2 \\ \mathbf{x}_k^3 \end{bmatrix}, \quad \mathbf{x}_k^1 = \begin{bmatrix} \boldsymbol{\Sigma}_k \\ \dot{\boldsymbol{\Sigma}}_k \\ \ddot{\boldsymbol{\Sigma}}_k \end{bmatrix}, \quad \mathbf{x}_k^2 = \begin{bmatrix} \Delta A_k \\ \Delta \dot{A}_k \\ \Delta \ddot{A}_k \end{bmatrix}, \quad \mathbf{x}_k^3 = \begin{bmatrix} \Delta E_k \\ \Delta \dot{E}_k \\ \Delta \ddot{E}_k \end{bmatrix} \\ \underline{\xi}_k &= \begin{bmatrix} \underline{\xi}_k^1 \\ \underline{\xi}_k^2 \\ \underline{\xi}_k^3 \end{bmatrix}, \quad \underline{\eta}_k = \begin{bmatrix} \underline{\eta}_k^1 \\ \underline{\eta}_k^2 \\ \underline{\eta}_k^3 \end{bmatrix}, \quad \mathbf{v}_k = \begin{bmatrix} \mathbf{v}_k^1 \\ \mathbf{v}_k^2 \\ \mathbf{v}_k^3 \end{bmatrix} \end{aligned}$$

$$\mathbf{A} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = [1 \ 0 \ 0]$$

由练习 3.8 所描述的系统可以分解为 3 个子系统：

$$\begin{cases} \mathbf{x}_{k+1}^i = \mathbf{Ax}_k^i + \boldsymbol{\Gamma}_k^i \underline{\boldsymbol{\xi}}_k^i \\ \mathbf{v}_k^i = \mathbf{Cx}_k^i + \eta_k^i \end{cases}$$

$i=1, 2, 3$ 。

对于每个 k , \mathbf{x}_k 和 $\underline{\boldsymbol{\xi}}_k$ 是 3 维向量, 相应的 v_k 和 η_k 是标量, \mathbf{Q}_k 是 3×3 的非负正定对称矩阵, $R_k > 0$ 也是标量。

第 4 章

4.1 由式(4.6), 有

$$\begin{aligned} L(\mathbf{Ax} + \mathbf{By}, \mathbf{v}) &= E(\mathbf{Ax} + \mathbf{By}) + \langle \mathbf{Ax} + \mathbf{By}, \mathbf{v} \rangle [\text{Var}(\mathbf{v})]^{-1} (\mathbf{v} - E(\mathbf{v})) \\ &= \mathbf{A}\{E(\mathbf{x}) + \langle \mathbf{x}, \mathbf{v} \rangle [\text{Var}(\mathbf{v})]^{-1} (\mathbf{v} - E(\mathbf{v}))\} \\ &\quad + \mathbf{B}\{E(\mathbf{y}) + \langle \mathbf{y}, \mathbf{v} \rangle [\text{Var}(\mathbf{v})]^{-1} (\mathbf{v} - E(\mathbf{v}))\} \\ &= \mathbf{AL}(\mathbf{x}, \mathbf{v}) + \mathbf{BL}(\mathbf{y}, \mathbf{v}) \end{aligned}$$

4.2 由式(4.6)及 $E(\mathbf{a}) = \mathbf{a}$, 有

$$\langle \mathbf{a}, \mathbf{v} \rangle = E(\mathbf{a} - E(\mathbf{a}))(\mathbf{v} - E(\mathbf{v})) = \mathbf{0}$$

从而

$$L(\mathbf{a}, \mathbf{v}) = E(\mathbf{a}) + \langle \mathbf{a}, \mathbf{v} \rangle [\text{Var}(\mathbf{v})]^{-1} (\mathbf{v} - E(\mathbf{v})) = \mathbf{a}$$

4.3 由定义, 对于实值函数 f 和矩阵 $\mathbf{A} = [a_{ij}]$, $d\mathbf{f}/d\mathbf{A} = [\partial f / \partial a_{ij}]$, 则

$$\begin{aligned} \mathbf{0} &= \frac{\partial}{\partial \mathbf{H}} (\text{tr} \parallel \mathbf{x} - \mathbf{y} \parallel_q^2) \\ &= \frac{\partial}{\partial \mathbf{H}} E((\mathbf{x} - E(\mathbf{x})) - \mathbf{H}(\mathbf{v} - E(\mathbf{v})))^\top ((\mathbf{x} - E(\mathbf{x})) - \mathbf{H}(\mathbf{v} - E(\mathbf{v}))) \\ &= E \frac{\partial}{\partial \mathbf{H}} ((\mathbf{x} - E(\mathbf{x})) - \mathbf{H}(\mathbf{v} - E(\mathbf{v})))^\top ((\mathbf{x} - E(\mathbf{x})) - \mathbf{H}(\mathbf{v} - E(\mathbf{v}))) \\ &= E(-2(\mathbf{x} - E(\mathbf{x})) - \mathbf{H}(\mathbf{v} - E(\mathbf{v}))) (\mathbf{v} - E(\mathbf{v}))^\top \\ &= 2(\mathbf{H} E(\mathbf{v} - E(\mathbf{v})) (\mathbf{v} - E(\mathbf{v}))^\top - E(\mathbf{x} - E(\mathbf{x})) (\mathbf{v} - E(\mathbf{v}))^\top) \\ &= 2(\mathbf{H} \parallel \mathbf{v} \parallel_q^2 - \langle \mathbf{x}, \mathbf{v} \rangle) \end{aligned}$$

这就可以给出

$$\mathbf{H}^* = \langle \mathbf{x}, \mathbf{v} \rangle [\parallel \mathbf{v} \parallel_q^2]^{-1}$$

使得

$$\mathbf{x}^* = E(\mathbf{x}) - \langle \mathbf{x}, \mathbf{v} \rangle [\parallel \mathbf{v} \parallel_q^2]^{-1} (E(\mathbf{v}) - \mathbf{v})$$

4.4 由于 \mathbf{v}^{k-2} 是下面各项的线性组合(常值矩阵系数):

$$x_0, \underline{\xi}_0, \dots, \underline{\xi}_{k-3}, \underline{\eta}_0, \dots, \underline{\eta}_{k-2}$$

这些都跟 $\underline{\xi}_{k-1}$ 和 $\underline{\eta}_{k-1}$ 不相关, 于是有

$$\langle \underline{\xi}_{k-1}, v^{k-2} \rangle = 0 \quad \text{和} \quad \langle \underline{\eta}_{k-1}, v^{k-2} \rangle = 0$$

同理, 可以证明其他的公式[可能会用到式(4.6)]。

4.5 第一个等式可以从卡尔曼增益方程得到(见定理4.1(c)或式(4.19)), 即 $\mathbf{G}_k (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k) = \mathbf{P}_{k,k-1} \mathbf{C}_k^T$, 使得 $\mathbf{G}_k \mathbf{R}_k = \mathbf{P}_{k,k-1} \mathbf{C}_k^T - \mathbf{G}_k \mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k,k-1} \mathbf{C}_k^T$ 。

为了证明第二个等式, 应用式(4.18)和式(4.17), 得到

$$\begin{aligned} & \langle x_{k-1} - \hat{x}_{k-1|k-1}, \Gamma_{k-1} \underline{\xi}_{k-1} - K_{k-1} \underline{\eta}_{k-1} \rangle \\ &= \langle x_{k-1} - \hat{x}_{k-1|k-2} - \langle x_{k-1}^\# , v_{k-1}^\# \rangle [\| v_{k-1}^\# \|^2]^{-1} v_{k-1}^\# , \Gamma_{k-1} \underline{\xi}_{k-1} - K_{k-1} \underline{\eta}_{k-1} \rangle \\ &= \langle x_{k-1}^\# - \langle x_{k-1}^\# , v_{k-1}^\# \rangle [\| v_{k-1}^\# \|^2]^{-1} (\mathbf{C}_{k-1} x_{k-1}^\# + \underline{\eta}_{k-1}) , \Gamma_{k-1} \underline{\xi}_{k-1} - K_{k-1} \underline{\eta}_{k-1} \rangle \\ &= -\langle x_{k-1}^\# , v_{k-1}^\# \rangle [\| v_{k-1}^\# \|^2]^{-1} (\mathbf{S}_{k-1}^T \Gamma_{k-1}^T - \mathbf{R}_{k-1} K_{k-1}^T) \\ &= \mathbf{O}_{n \times n} \end{aligned}$$

又 $K_{k-1} = \Gamma_{k-1} S_{k-1} R_{k-1}^{-1}$, 可得

$$\mathbf{S}_{k-1}^T \Gamma_{k-1}^T - \mathbf{R}_{k-1} K_{k-1}^T = \mathbf{O}_{n \times n}$$

4.6 与推导定理4.1的过程一样, 用 $v_k - D_k u_k$ 代替 v_k , 并且用

$$\hat{x}_{k|k-1} = L(\mathbf{A}_{k-1} x_{k-1} + \mathbf{B}_{k-1} u_{k-1} + \Gamma_{k-1} \underline{\xi}_{k-1}, v^{k-1})$$

代替

$$\hat{x}_{k|k-1} = L(x_k, v^{k-1}) = L(\mathbf{A}_{k-1} x_{k-1} + \Gamma_{k-1} \underline{\xi}_{k-1}, v^{k-1})$$

4.7 设

$$w_k = -a_1 v_{k-1} + b_1 u_{k-1} + c_1 e_{k-1} + w_{k-1}$$

$$w_{k-1} = -a_2 v_{k-2} + b_2 u_{k-2} + w_{k-2}$$

$$w_{k-2} = -a_3 v_{k-3}$$

并定义 $x_k = [w_k \ w_{k-1} \ w_{k-2}]^T$, 则

$$\begin{cases} x_{k+1} = \mathbf{A} x_k + \mathbf{B} u_k + \boldsymbol{\Gamma} e_k \\ v_k = \mathbf{C} x_k + \mathbf{D} u_k + \Delta e_k \end{cases}$$

式中

$$\mathbf{A} = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 - a_1 b_0 \\ b_2 - a_2 b_0 \\ -a_3 b_0 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} c_1 - a_1 c_0 \\ -a_2 c_0 \\ -a_3 c_0 \end{bmatrix}$$

$$\mathbf{C} = [1 \ 0 \ 0], \quad \mathbf{D} = [b_0], \quad \Delta = [c_0]$$

4.8 设

$$\begin{aligned} w_k &= -a_1 v_{k-1} + b_1 u_{k-1} + c_1 e_{k-1} + w_{k-1} \\ w_{k-1} &= -a_2 v_{k-2} + b_2 u_{k-2} + c_2 e_{k-2} + w_{k-2} \\ &\vdots \\ w_{k-n+1} &= -a_n v_{k-n} + b_n u_{k-n} + c_n e_{k-n} \end{aligned}$$

对于 $j > m$ 有 $b_j = 0$, 对于 $j > l$ 有 $c_j = 0$, 并且定义

$$\mathbf{x}_k = [w_k \quad w_{k-1} \quad \cdots \quad w_{k-n+1}]^T$$

则

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k + \boldsymbol{\Gamma}e_k \\ \mathbf{v}_k = \mathbf{C}\mathbf{x}_k + Du_k + \Delta e_k \end{cases}$$

式中

$$\mathbf{A} = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -a_{n-1} & 0 & 0 & \cdots & 1 \\ -a_n & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 - a_1 b_0 \\ \vdots \\ b_m - a_m b_0 \\ -a_{m+1} b_0 \\ \vdots \\ -a_n b_0 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} c_1 - a_1 c_0 \\ \vdots \\ c_l - a_l c_0 \\ -a_{l+1} \\ \vdots \\ -a_n c_0 \end{bmatrix}$$

$$\mathbf{C} = [1 \quad 0 \quad \cdots \quad 0], \quad D = [b_0], \quad \Delta = [c_0]$$

第 5 章

5.1 \mathbf{v}^k 是下面各项的线性组合(常值矩阵系数):

$$\mathbf{x}_0, \underline{\gamma}_0, \underline{\gamma}_0, \cdots, \underline{\gamma}_k, \underline{\xi}_0, \underline{\beta}_0, \cdots, \underline{\beta}_{k-1}$$

所有这些都独立于 $\underline{\beta}_k$, 有 $\langle \underline{\beta}_k, \mathbf{v}^k \rangle = \mathbf{0}$ 。另一方面, $\underline{\beta}_k$ 是零均值的。由式(4.6), 有

$$L(\hat{\underline{\beta}}_k, \mathbf{v}^k) = E(\hat{\underline{\beta}}_k) - \langle \hat{\underline{\beta}}_k, \mathbf{v}^k \rangle [\|\mathbf{v}^k\|^2]^{-1} (E(\mathbf{v}^k) - \mathbf{v}^k) = \mathbf{0}$$

5.2 由引理 4.2, 取 $\mathbf{v} = \mathbf{v}^{k-1}$, $\mathbf{v}^1 = \mathbf{v}^{k-2}$, $\mathbf{v}^2 = \mathbf{v}_{k-1}$, $\mathbf{v}_{k-1}^\# = \mathbf{v}_{k-1} - L(\mathbf{v}_{k-1}, \mathbf{v}^{k-2})$, 则对于所有 $\mathbf{x} = \mathbf{v}_{k-1}$, 有

$$\begin{aligned} L(\mathbf{v}_{k-1}, \mathbf{v}^{k-1}) &= L(\mathbf{v}_{k-1}, \mathbf{v}^{k-2}) + \langle \mathbf{v}_{k-1}^\#, \mathbf{v}_{k-1}^\# \rangle [\|\mathbf{v}_{k-1}^\#\|^2]^{-1} \mathbf{v}_{k-1}^\# \\ &= L(\mathbf{v}_{k-1}, \mathbf{v}^{k-2}) + \mathbf{v}_{k-1} - L(\mathbf{v}_{k-1}, \mathbf{v}^{k-2}) = \mathbf{v}_{k-1} \end{aligned}$$

等式 $L(\underline{\gamma}_k, \mathbf{v}^{k-1}) = \mathbf{0}$, 可以模仿练习 5.1 的证明来完成。

5.3 根据引理 4.2 有

$$\begin{aligned} \mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1} &= \mathbf{z}_{k-1} - L(\mathbf{z}_{k-1}, \mathbf{v}^{k-1}) \\ &= \mathbf{z}_{k-1} - E(\mathbf{z}_{k-1}) + \langle \mathbf{z}_{k-1}, \mathbf{v}^{k-1} \rangle [\|\mathbf{v}^{k-1}\|^2]^{-1} (E(\mathbf{v}^{k-1}) - \mathbf{v}^{k-1}) \end{aligned}$$

$$= \begin{bmatrix} \mathbf{x}_{k-1} \\ \underline{\xi}_{k-1} \end{bmatrix} - \begin{bmatrix} E(\mathbf{x}_{k-1}) \\ E(\underline{\xi}_{k-1}) \end{bmatrix} + \begin{bmatrix} \langle \mathbf{x}_{k-1}, \mathbf{v}^{k-1} \rangle \\ \langle \underline{\xi}_{k-1}, \mathbf{v}^{k-1} \rangle \end{bmatrix} [\|\mathbf{v}^{k-1}\|^2]^{-1} (E(\mathbf{v}^{k-1}) - \mathbf{v}^{k-1})$$

其最前 n 维子向量和最后 p 维子向量分别是下面的线性组合(以常值矩阵为系数):

$$\mathbf{x}_0, \underline{\xi}_0, \underline{\beta}_0, \dots, \underline{\beta}_{k-2}, \underline{\eta}_0, \underline{\gamma}_0, \dots, \underline{\gamma}_{k-1}$$

它们都独立于 $\underline{\gamma}_k$, 因此有

$$B(\mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}, \underline{\gamma}_k) = \mathbf{0}$$

5.4 该证明和练习 5.3 类似。

5.5 为了简化, 记

$$\mathbf{B} = [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T + \mathbf{R}_0]^{-1}$$

由式(5.16), 有

$$\begin{aligned} \text{Var}(\mathbf{x}_0 - \hat{\mathbf{x}}_0) &= \text{Var}(\mathbf{x}_0 - E(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T + \mathbf{R}_0]^{-1} (\mathbf{v}_0 - \mathbf{C}_0 E(\mathbf{x}_0))) \\ &= \text{Var}(\mathbf{x}_0 - E(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} (\mathbf{C}_0 (\mathbf{x}_0 - E(\mathbf{x}_0)) + \underline{\eta}_0)) \\ &= \text{Var}((\mathbf{I} - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0) (\mathbf{x}_0 - E(\mathbf{x}_0)) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \underline{\eta}_0) \\ &= (\mathbf{I} - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0) \text{Var}(\mathbf{x}_0) (\mathbf{I} - \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)]) \\ &\quad + [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{R}_0 \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \\ &= \text{Var}(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \\ &\quad + [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \\ &\quad + [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{R}_0 \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \\ &= \text{Var}(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \\ &\quad + [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \\ &= \text{Var}(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T \mathbf{B} \mathbf{C}_0 [\text{Var}(\mathbf{x}_0)] \end{aligned}$$

5.6 由 $\hat{\xi}_0 = \mathbf{0}$, 有 $\hat{\mathbf{x}}_1 = \mathbf{A}_0 \hat{\mathbf{x}}_0 + \mathbf{G}_1 (\mathbf{v}_1 - \mathbf{C}_1 \mathbf{A}_0 \hat{\mathbf{x}}_0)$, $\hat{\xi}_1 = \mathbf{0}$, 从而有 $\hat{\mathbf{x}}_2 = \mathbf{A}_1 \hat{\mathbf{x}}_1 + \mathbf{G}_2 (\mathbf{v}_2 - \mathbf{C}_2 \mathbf{A}_1 \hat{\mathbf{x}}_1)$, 等等。

一般地, 有

$$\hat{\mathbf{x}}_k = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1}) = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})$$

定义

$$\mathbf{P}_{0,0} = [[\text{Var}(\mathbf{x}_0)]^{-1} + \mathbf{C}_0^T \mathbf{R}_0^{-1} \mathbf{C}_0]^{-1}$$

$$\mathbf{P}_{k,k-1} = \mathbf{A}_{k-1} \mathbf{P}_{k-1,k-1} \mathbf{A}_{k-1}^T + \mathbf{\Gamma}_{k-1} \mathbf{Q}_{k-1} \mathbf{\Gamma}_{k-1}^T$$

则

$$\mathbf{G}_1 = \begin{bmatrix} \mathbf{A}_0 & \mathbf{\Gamma}_0 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{0,0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix} \begin{bmatrix} \mathbf{A}_0^T & \mathbf{C}_1^T \\ \mathbf{\Gamma}_0^T & \mathbf{C}_1^T \end{bmatrix} \cdot \left([\mathbf{C}_1 \mathbf{A}_0 \quad \mathbf{C}_1 \mathbf{\Gamma}_0] \begin{bmatrix} \mathbf{P}_{0,0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix} \begin{bmatrix} \mathbf{A}_0^T & \mathbf{C}_1^T \\ \mathbf{\Gamma}_0^T & \mathbf{C}_1^T \end{bmatrix} + \mathbf{R}_1 \right)^{-1}$$

$$\begin{aligned}
&= \begin{bmatrix} \mathbf{P}_{1,0} \mathbf{C}_1^T (\mathbf{C}_1 \mathbf{P}_{1,0} \mathbf{C}_1^T + \mathbf{R}_1)^{-1} \\ \mathbf{0} \end{bmatrix} \\
\mathbf{P}_1 &= \left(\begin{bmatrix} \mathbf{A}_0 & \boldsymbol{\Gamma}_0 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} - \mathbf{G}_1 [\mathbf{C}_1 \mathbf{A}_0 \quad \mathbf{C}_1 \boldsymbol{\Gamma}_0] \right) \begin{bmatrix} \mathbf{P}_{0,0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix} \begin{bmatrix} \mathbf{A}_0^T & \mathbf{0} \\ \boldsymbol{\Gamma}_0^T & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix} \\
&= \begin{bmatrix} [\mathbf{I}_n - \mathbf{P}_{1,0} \mathbf{C}_1^T (\mathbf{C}_1 \mathbf{P}_{1,0} \mathbf{C}_1^T + \mathbf{R}_1)^{-1} \mathbf{C}_1] \mathbf{P}_{1,0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix}
\end{aligned}$$

一般地,有

$$\begin{aligned}
\mathbf{G}_k &= \begin{bmatrix} \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ \mathbf{0} \end{bmatrix} \\
\mathbf{P}_k &= \begin{bmatrix} [\mathbf{I}_n - \mathbf{P}_{k,k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k,k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \mathbf{C}_k] \mathbf{P}_{k,k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_k \end{bmatrix}
\end{aligned}$$

最后,用 \mathbf{x}_0 的无偏估计 $\hat{\mathbf{x}}_0 = E(\mathbf{x}_0)$ 代替稍微好一些的初始状态估计:

$$\hat{\mathbf{x}}_0 = E(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^T [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^T + \mathbf{R}_0]^{-1} [\mathbf{C}_0 E(\mathbf{x}_0) - \mathbf{v}_0]$$

并设

$$\mathbf{P}_0 = E \left(\begin{bmatrix} \mathbf{x}_0 \\ \underline{\xi}_0 \end{bmatrix} - \begin{bmatrix} E(\mathbf{x}_0) \\ E(\underline{\xi}_0) \end{bmatrix} \right) \left(\begin{bmatrix} \mathbf{x}_0 \\ \underline{\xi}_0 \end{bmatrix} - \begin{bmatrix} E(\mathbf{x}_0) \\ E(\underline{\xi}_0) \end{bmatrix} \right)^T = \begin{bmatrix} \text{Var}(\mathbf{x}_0) & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix}$$

则可以得到第 2 章和第 3 章的卡尔曼滤波算法。

5.7 设

$$\begin{aligned}
\bar{\mathbf{P}}_0 &= [[\text{Var}(\mathbf{x}_0)]^{-1} + \mathbf{C}_0^T \mathbf{R}_0^{-1} \mathbf{C}_0]^{-1} \\
\bar{\mathbf{H}}_{k-1} &= [\mathbf{C}_k \mathbf{A}_{k-1} - \mathbf{N}_{k-1} \mathbf{C}_{k-1}]
\end{aligned}$$

从(5. 17b)开始,即

$$\mathbf{P}_0 = \begin{bmatrix} ([\text{Var}(\mathbf{x}_0)]^{-1} + \mathbf{C}_0 \mathbf{R}_0^{-1} \mathbf{C}_0)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{P}}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix}$$

有

$$\begin{aligned}
\mathbf{G}_1 &= \begin{bmatrix} \mathbf{A}_0 & \boldsymbol{\Gamma}_0 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{P}}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix} \begin{bmatrix} \bar{\mathbf{H}}_0^T \\ \boldsymbol{\Gamma}_0^T \mathbf{C}_1^T \end{bmatrix} \cdot \begin{bmatrix} [\mathbf{H}_0 \quad \mathbf{C}_1 \boldsymbol{\Gamma}_0] \end{bmatrix} \begin{bmatrix} \bar{\mathbf{P}}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix} \begin{bmatrix} \bar{\mathbf{H}}_0^T \\ \boldsymbol{\Gamma}_0^T \mathbf{C}_1^T \end{bmatrix} + \mathbf{R}_1 \Big]^{-1} \\
&= \begin{bmatrix} (\mathbf{A}_0 \bar{\mathbf{P}}_0 \bar{\mathbf{H}}_0^T + \boldsymbol{\Gamma}_0 \mathbf{Q}_0 \boldsymbol{\Gamma}_0^T \mathbf{C}_1^T) (\bar{\mathbf{H}}_0 \bar{\mathbf{P}}_0 \bar{\mathbf{H}}_0^T + \mathbf{C}_1 \mathbf{Q}_0 \boldsymbol{\Gamma}_0^T \mathbf{C}_1^T + \mathbf{R}_1)^{-1} \\ \mathbf{0} \end{bmatrix} := \begin{bmatrix} \bar{\mathbf{G}}_1 \\ \mathbf{0} \end{bmatrix} \\
\mathbf{P}_1 &= \left[\begin{bmatrix} \mathbf{A}_0 & \boldsymbol{\Gamma}_0 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \bar{\mathbf{G}}_1 \\ \mathbf{0} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{H}}_0 & \mathbf{C}_1 \boldsymbol{\Gamma}_0 \end{bmatrix} \right] \begin{bmatrix} \bar{\mathbf{P}}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0 \end{bmatrix} \begin{bmatrix} \mathbf{A}_0^T & \mathbf{0} \\ \boldsymbol{\Gamma}_0^T & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix}
\end{aligned}$$

$$= \begin{bmatrix} (\mathbf{A}_0 - \bar{\mathbf{G}}_1 \bar{\mathbf{H}}_0) \bar{\mathbf{P}}_0 \mathbf{A}_0^\top + (I - \bar{\mathbf{G}}_1 \mathbf{C}_1) \boldsymbol{\Gamma}_0 \mathbf{Q}_0 \boldsymbol{\Gamma}_0^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix} := \begin{bmatrix} \bar{\mathbf{P}}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix}$$

一般地, 可得

$$\left\{ \begin{array}{l} \hat{\mathbf{x}}_k = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1} + \bar{\mathbf{G}}_k (\mathbf{v}_k - \mathbf{N}_{k-1} \mathbf{v}_{k-1} - \bar{\mathbf{H}}_{k-1} \hat{\mathbf{x}}_{k-1}) \\ \hat{\mathbf{x}}_0 = E(\mathbf{x}_0) - [\text{Var}(\mathbf{x}_0)] \mathbf{C}_0^\top [\mathbf{C}_0 \text{Var}(\mathbf{x}_0) \mathbf{C}_0^\top + \mathbf{R}_0]^{-1} [\mathbf{C}_0 E(\mathbf{x}_0) - \mathbf{v}_0] \\ \bar{\mathbf{H}}_{k-1} = [\mathbf{C}_k \mathbf{A}_{k-1} - \mathbf{N}_{k-1} \mathbf{C}_{k-1}] \\ \bar{\mathbf{P}}_k = (\mathbf{A}_{k-1} - \bar{\mathbf{G}}_k \bar{\mathbf{H}}_{k-1}) \bar{\mathbf{P}}_{k-1} \mathbf{A}_{k-1}^\top + (I - \bar{\mathbf{G}}_k \mathbf{C}_k) \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^\top \\ \bar{\mathbf{G}}_k = (\mathbf{A}_{k-1} \bar{\mathbf{P}}_{k-1} \bar{\mathbf{H}}_{k-1}^\top + \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^\top \mathbf{C}_k^\top) \cdot (\bar{\mathbf{H}}_{k-1} \bar{\mathbf{P}}_{k-1} \bar{\mathbf{H}}_{k-1}^\top \\ \quad + \mathbf{C}_k \boldsymbol{\Gamma}_{k-1} \mathbf{Q}_{k-1} \boldsymbol{\Gamma}_{k-1}^\top \mathbf{C}_k^\top + \mathbf{R}_{k-1})^{-1} \\ \bar{\mathbf{P}}_0 = [[\text{Var}(\mathbf{x}_0)]^{-1} + \mathbf{C}_0^\top \mathbf{R}_0^{-1} \mathbf{C}_0]^{-1} \\ k = 1, 2, \dots \end{array} \right.$$

通过忽略 $\bar{\mathbf{H}}_k$ 、 $\bar{\mathbf{G}}_k$ 和 $\bar{\mathbf{P}}_k$ 上的“横线”, 就可以得到式(5.21)。

$$5.8 \quad (a) \quad \begin{cases} \underline{\mathbf{X}}_{k+1} = \mathbf{A}_c \underline{\mathbf{X}}_k + \underline{\boldsymbol{\xi}}_k \\ \mathbf{v}_k = \mathbf{C}_c \underline{\mathbf{X}}_k \end{cases}$$

$$(b) \quad \mathbf{P}_{0,0} = \begin{bmatrix} \text{Var}(\mathbf{x}_0) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \text{Var}(\underline{\boldsymbol{\xi}}_0) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \text{Var}(\eta_0) \end{bmatrix}$$

$$\mathbf{P}_{k,k-1} = \mathbf{A}_c \mathbf{P}_{k-1,k-1} \mathbf{A}_c^\top + \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & \mathbf{Q}_{k-1} \\ & & & r_{k-1} \end{bmatrix}$$

$$\mathbf{G}_k = \mathbf{P}_{k,k-1} \mathbf{C}_c^\top (\mathbf{C}_c^\top \mathbf{P}_{k,k-1} \mathbf{C}_c)^{-1}$$

$$\mathbf{P}_{k,k} = (I - \mathbf{G}_k \mathbf{C}_c) \mathbf{P}_{k,k-1}$$

$$\underline{\hat{\mathbf{X}}}_0 = \begin{bmatrix} E(\mathbf{x}_0) \\ \mathbf{0} \\ 0 \end{bmatrix}$$

$$\hat{\mathbf{x}}_k = \mathbf{A}_c \underline{\hat{\mathbf{X}}}_{k-1} + \mathbf{G}_k (\mathbf{v}_k - \mathbf{C}_c \mathbf{A}_c \underline{\hat{\mathbf{X}}}_{k-1})$$

(c) 矩阵 $\mathbf{C}_c^\top \mathbf{P}_{k,k-1} \mathbf{C}_c$ 可能不可逆, 需要额外估计 $\underline{\hat{\mathbf{X}}}_k$ 中的 $\underline{\boldsymbol{\xi}}_k$ 和 $\hat{\boldsymbol{\eta}}_k$ 。

第 6 章

6.1 由于

$$\begin{aligned}\mathbf{x}_{k-1} &= \mathbf{A}\mathbf{x}_{k-2} + \boldsymbol{\Gamma}\boldsymbol{\xi}_{k-2} = \cdots = \mathbf{A}^n\mathbf{x}_{k-n-1} + \text{噪声} \\ \tilde{\mathbf{x}}_{k-1} &= \mathbf{A}^n[\mathbf{N}_{CA}^T \mathbf{N}_{CA}]^{-1}(\mathbf{C}^T \mathbf{v}_{k-n-1} + \mathbf{A}^T \mathbf{C}^T \mathbf{v}_{k-n} + \cdots + (\mathbf{A}^T)^{n-1} \mathbf{C}^T \mathbf{v}_{k-2}) \\ &= \mathbf{A}^n[\mathbf{N}_{CA}^T \mathbf{N}_{CA}]^{-1}(\mathbf{C}^T \mathbf{C} \mathbf{x}_{k-n-1} + \mathbf{A}^T \mathbf{C}^T \mathbf{C} \mathbf{A} \mathbf{x}_{k-n-1} + \cdots \\ &\quad + (\mathbf{A}^T)^{n-1} \mathbf{C}^T \mathbf{C} \mathbf{A}^{n-1} \mathbf{x}_{k-n-1} + \text{噪声}) \\ &= \mathbf{A}^n[\mathbf{N}_{CA}^T \mathbf{N}_{CA}]^{-1}[\mathbf{N}_{CA}^T \mathbf{N}_{CA}] \mathbf{x}_{k-n-1} + \text{噪声} \\ &= \mathbf{A}^n \mathbf{x}_{k-n-1} + \text{噪声}\end{aligned}$$

有

$$E(\tilde{\mathbf{x}}_{k-1}) = E(\mathbf{A}^n \mathbf{x}_{k-n-1}) = E(\mathbf{x}_{k-1})$$

6.2 由于

$$\frac{d}{ds}[\mathbf{A}^{-1}(s)\mathbf{A}(s)] = \frac{d}{ds}\mathbf{I} = \mathbf{0}$$

有

$$\mathbf{A}^{-1}(s) \left[\frac{d}{ds} \mathbf{A}(s) \right] + \left[\frac{d}{ds} \mathbf{A}^{-1}(s) \right] \mathbf{A}(s) = \mathbf{0}$$

从而

$$\frac{d}{ds} \mathbf{A}^{-1}(s) = -\mathbf{A}^{-1}(s) \left[\frac{d}{ds} \mathbf{A}(s) \right] \mathbf{A}^{-1}(s)$$

6.3 设 $\mathbf{P} = \mathbf{U} \text{diag}[\lambda_1, \dots, \lambda_n] \mathbf{U}^{-1}$, 则

$$\mathbf{P} - \lambda_{\min} \mathbf{I} = \mathbf{U} \text{diag}[\lambda_1 - \lambda_{\min}, \dots, \lambda_n - \lambda_{\min}] \mathbf{U}^{-1} \geq \mathbf{0}.$$

6.4 设 $\lambda_1, \dots, \lambda_n$ 是 \mathbf{F} 的特征值, \mathbf{J} 是其若尔当标准型, 则存在非奇异矩阵 \mathbf{U} , 使得

$$\mathbf{U}^{-1} \mathbf{F} \mathbf{U} = \mathbf{J} = \begin{bmatrix} \lambda_1 & * & & & \\ & \lambda_2 & * & & \\ & & \ddots & \ddots & \\ & & & \ddots & * \\ & & & & \lambda_n \end{bmatrix}$$

式中每个 * 为 1 或 0。因此

$$\mathbf{F}^k = \mathbf{U} \mathbf{J}^k \mathbf{U}^{-1} = \begin{bmatrix} \lambda_1^k & * & \cdots & \cdots & * \\ & \lambda_2^k & * & \cdots & * \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_n^k \end{bmatrix}$$

式中每个 * 表示以下式为界的项: $p(k) |\lambda_{\max}|^k$, 其中 $p(k)$ 是 k 的多项式, $|\lambda_{\max}| =$

$\max(|\lambda_1|, \dots, |\lambda_n|)$ 。

因为 $|\lambda_{\max}| < 1$, 当 $k \rightarrow \infty$ 时, $\mathbf{F}^k \rightarrow 0$ 。

6.5 因为

$$\mathbf{0} \leq (\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})^T = \mathbf{AA}^T - \mathbf{AB}^T - \mathbf{BA}^T + \mathbf{BB}^T$$

有

$$\mathbf{AB}^T + \mathbf{BA}^T \leq \mathbf{AA}^T + \mathbf{BB}^T$$

因此

$$(\mathbf{A} + \mathbf{B})(\mathbf{A} + \mathbf{B})^T = \mathbf{AA}^T + \mathbf{AB}^T + \mathbf{BA}^T + \mathbf{BB}^T \leq 2(\mathbf{AA}^T + \mathbf{BB}^T)$$

6.6 由于 $\mathbf{x}_{k-1} = \mathbf{Ax}_{k-2} + \boldsymbol{\Gamma}\underline{\xi}_{k-2}$ 是 $\mathbf{x}_0, \underline{\xi}_0, \dots, \underline{\xi}_{k-2}$ 的线性组合(以常值矩阵为系数), 则 $\vec{\mathbf{x}}_{k-1} = \mathbf{A}\vec{\mathbf{x}}_{k-2} + \mathbf{G}(\mathbf{v}_{k-1} - \mathbf{CA}\vec{\mathbf{x}}_{k-2}) = \mathbf{A}\vec{\mathbf{x}}_{k-2} + \mathbf{G}(\mathbf{CAx}_{k-2} + \boldsymbol{\Gamma}\underline{\xi}_{k-2} + \underline{\eta}_{k-1}) - \mathbf{GCA}\vec{\mathbf{x}}_{k-2}$ 是 $\mathbf{x}_0, \underline{\xi}_0, \dots, \underline{\xi}_{k-2}$ 和 $\underline{\eta}_{k-1}$ 的类似线性组合, 再由 $\underline{\xi}_{k-1}, \underline{\eta}_k$ 不相关, 这两个恒等式就可以得到了。

6.7 因为

$$\begin{aligned} \mathbf{P}_{k,k-1}\mathbf{C}_k^T\mathbf{G}_k^T - \mathbf{G}_k\mathbf{C}_k\mathbf{P}_{k,k-1}\mathbf{C}_k^T\mathbf{G}_k^T &= \mathbf{G}_k\mathbf{C}_k\mathbf{P}_{k,k-1}\mathbf{C}_k^T\mathbf{G}_k^T + \mathbf{G}_k\mathbf{R}_k\mathbf{G}_k^T - \mathbf{G}_k\mathbf{C}_k\mathbf{P}_{k,k-1}\mathbf{C}_k^T\mathbf{G}_k^T \\ &= \mathbf{G}_k\mathbf{R}_k\mathbf{G}_k^T, \end{aligned}$$

有

$$-(\mathbf{I} - \mathbf{G}_k\mathbf{C})\mathbf{P}_{k,k-1}\mathbf{C}^T\mathbf{G}_k^T + \mathbf{G}_k\mathbf{R}\mathbf{G}_k^T = \mathbf{0}$$

因此

$$\begin{aligned} \mathbf{P}_{k,k} &= (\mathbf{I} - \mathbf{G}_k\mathbf{C})\mathbf{P}_{k,k-1} \\ &= (\mathbf{I} - \mathbf{G}_k\mathbf{C})\mathbf{P}_{k,k-1}(\mathbf{I} - \mathbf{G}_k\mathbf{C})^T + \mathbf{G}_k\mathbf{R}\mathbf{G}_k^T \\ &= (\mathbf{I} - \mathbf{G}_k\mathbf{C})(\mathbf{A}\mathbf{P}_{k-1,k-1}\mathbf{A}^T + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T)(\mathbf{I} - \mathbf{G}_k\mathbf{C})^T + \mathbf{G}_k\mathbf{R}\mathbf{G}_k^T \\ &= (\mathbf{I} - \mathbf{G}_k\mathbf{C})\mathbf{A}\mathbf{P}_{k-1,k-1}\mathbf{A}^T(\mathbf{I} - \mathbf{G}_k\mathbf{C})^T + (\mathbf{I} - \mathbf{G}_k\mathbf{C})\boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^T(\mathbf{I} - \mathbf{G}_k\mathbf{C})^T + \mathbf{G}_k\mathbf{R}\mathbf{G}_k^T \end{aligned}$$

6.8 参照引理 6.8 的证明, 假设 $|\lambda| \geq 1$, 其中 λ 是 $(\mathbf{I} - \mathbf{GC})\mathbf{A}$ 的特征值, 可以得到与可控性条件矛盾的结论。

6.9 与练习 6.6 的证明类似。

6.10 由

$$\begin{aligned} \mathbf{0} &\leq \langle \underline{\epsilon}_j - \underline{\delta}_j, \underline{\epsilon}_j - \underline{\delta}_j \rangle \\ &= \langle \underline{\epsilon}_j, \underline{\epsilon}_j \rangle - \langle \underline{\epsilon}_j, \underline{\delta}_j \rangle - \langle \underline{\delta}_j, \underline{\epsilon}_j \rangle + \langle \underline{\delta}_j, \underline{\delta}_j \rangle \end{aligned}$$

和定理 6.2, 当 $j \rightarrow \infty$ 时, 有

$$\begin{aligned} \langle \underline{\epsilon}_j, \underline{\delta}_j \rangle + \langle \underline{\delta}_j, \underline{\epsilon}_j \rangle &\leq \langle \underline{\epsilon}_j, \underline{\epsilon}_j \rangle + \langle \underline{\delta}_j, \underline{\delta}_j \rangle \\ &= \langle \hat{\mathbf{x}}_j - \mathbf{x}_j + \mathbf{x}_j - \vec{\mathbf{x}}_j, \hat{\mathbf{x}}_j - \mathbf{x}_j + \mathbf{x}_j - \vec{\mathbf{x}}_j \rangle + \|\mathbf{x}_j - \vec{\mathbf{x}}_j\|_n^2 \\ &= \|\mathbf{x}_j - \hat{\mathbf{x}}_j\|_n^2 + \langle \mathbf{x}_j - \vec{\mathbf{x}}_j, \hat{\mathbf{x}}_j - \mathbf{x}_j \rangle + \langle \hat{\mathbf{x}}_j - \mathbf{x}_j, \mathbf{x}_j - \vec{\mathbf{x}}_j \rangle + 2\|\mathbf{x}_j - \vec{\mathbf{x}}_j\|_n^2 \\ &\leq 2\|\mathbf{x}_j - \hat{\mathbf{x}}_j\|_n^2 + 3\|\mathbf{x}_j - \hat{\mathbf{x}}_j\|_n^2 \rightarrow 5(\mathbf{P}^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C})^{-1} \end{aligned}$$

因此, $\mathbf{B}_j = \langle \boldsymbol{\epsilon}_j, \boldsymbol{\delta}_j \rangle \mathbf{A}^T \mathbf{C}^T$ 各个分量都是一致有界的。

6.11 由引理 1.4、引理 1.6、引理 1.7 和引理 1.10 及定理 6.1, 并应用练习 6.10, 对于实数 $r_1, 0 < r_1 < 1$, 存在与 i 和 k 无关的正常数 C , 满足

$$\begin{aligned} & \text{tr}[\mathbf{F}\mathbf{B}_{k-1-i}(\mathbf{G}_{k-i} - \mathbf{G})^T + (\mathbf{G}_{k-i} - \mathbf{G})\mathbf{B}_{k-1-i}^T \mathbf{F}^T] \\ & \leq (n \text{tr} \mathbf{F} \mathbf{B}_{k-1-i} (\mathbf{G}_{k-i} - \mathbf{G})^T (\mathbf{G}_{k-i} - \mathbf{G}) \mathbf{B}_{k-1-i}^T \mathbf{F}^T)^{1/2} \\ & \quad + (n \text{tr} (\mathbf{G}_{k-i} - \mathbf{G}) \mathbf{B}_{k-1-i}^T \mathbf{F}^T \mathbf{F} \mathbf{B}_{k-1-i} (\mathbf{G}_{k-i} - \mathbf{G})^T)^{1/2} \\ & \leq (n \text{tr} \mathbf{F} \mathbf{F}^T \cdot \text{tr} \mathbf{B}_{k-1-i} \mathbf{B}_{k-1-i}^T \cdot \text{tr} (\mathbf{G}_{k-i} - \mathbf{G})^T (\mathbf{G}_{k-i} - \mathbf{G}))^{1/2} \\ & \quad + (n \text{tr} (\mathbf{G}_{k-i} - \mathbf{G}) (\mathbf{G}_{k-i} - \mathbf{G})^T \cdot \text{tr} \mathbf{B}_{k-1-i}^T \mathbf{B}_{k-1-i} \cdot \text{tr} \mathbf{F}^T \mathbf{F})^{1/2} \\ & = 2(n \text{tr} (\mathbf{G}_{k-i} - \mathbf{G}) (\mathbf{G}_{k-i} - \mathbf{G})^T \cdot \text{tr} \mathbf{B}_{k-1-i}^T \mathbf{B}_{k-1-i} \cdot \text{tr} \mathbf{F}^T \mathbf{F})^{1/2} \\ & \leq C_1 r_1^{k+1-i} \end{aligned}$$

6.12 求解 Riccati 方程(6.6), 即

$$c^2 p^2 + [(1-a^2)r - c^2 \gamma^2 q]p - rq\gamma^2 = 0$$

可得

$$p = \frac{1}{2c^2} \{ c^2 \gamma^2 q + (a^2 - 1)r + \sqrt{[(1-a^2)r - c^2 \gamma^2 q]^2 + 4c^2 \gamma^2 qr} \}$$

然后, 可以得到卡尔曼增益

$$g = pc / (c^2 p + r)$$

第 7 章

7.1 引理 7.1 的证明是构造性的。设 $\mathbf{A} = [a_{ij}]_{n \times n}$ 和 $\mathbf{A}^c = [l_{ij}]_{n \times n}$ 。由 $\mathbf{A} = \mathbf{A}^c (\mathbf{A}^c)^T$, 可得

$$\begin{aligned} a_{ii} &= \sum_{k=1}^i l_{ik}^2, \quad i = 1, 2, \dots, n \\ a_{ij} &= \sum_{k=1}^j l_{ik} l_{jk}, \quad j \neq i; \quad i, j = 1, 2, \dots, n \end{aligned}$$

可以证明

$$\begin{aligned} l_{ii} &= (a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2)^{1/2}, \quad i = 1, 2, \dots, n \\ l_{ij} &= (a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}) / l_{jj}, \quad j = 1, 2, \dots, i-1; \quad i = 2, 3, \dots, n \\ l_{ij} &= 0, \quad j = i+1, i+2, \dots, n; \quad i = 1, 2, \dots, n \end{aligned}$$

这就给出了下三角矩阵 \mathbf{A}^c , 该算法称为 Cholesky 分解。

由(标准)奇异值分解(SVD: singular value decomposition)算法来得到正交矩阵 \mathbf{U} , 使得

$$\mathbf{U} \text{diag}[s_1, \dots, s_r, 0, \dots, 0] \mathbf{U}^T = \mathbf{A} \mathbf{A}^T$$

式中, $1 \leq r \leq n$, s_1, \dots, s_r 是非负定对称矩阵 $\mathbf{A} \mathbf{A}^T$ 的奇异值(为正数)。设置

$$\tilde{\mathbf{A}} = \mathbf{U} \operatorname{diag}[\sqrt{s_1}, \sqrt{s_2}, \dots, \sqrt{s_r}, 0, \dots, 0]$$

7.2

$$(a) \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & -2 & 1 \end{bmatrix} \quad (b) \mathbf{L} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ \sqrt{2}/2 & \sqrt{2.5} & 0 \\ \sqrt{2}/2 & 1.5/\sqrt{2.5} & \sqrt{2.6} \end{bmatrix}$$

7.3

$$(a) \mathbf{L}^{-1} = \begin{bmatrix} 1/l_{11} & 0 & 0 \\ -l_{21}/l_{11}l_{22} & 1/l_{22} & 0 \\ -l_{31}/l_{11}l_{33} + l_{32}l_{21}/l_{11}l_{22}l_{33} & -l_{32}/l_{22}l_{33} & 1/l_{33} \end{bmatrix}$$

$$(b) \mathbf{L}^{-1} = \begin{bmatrix} b_{11} & 0 & 0 & \cdots & 0 \\ b_{21} & b_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & 0 \\ b_{n1} & b_{n2} & b_{n3} & \cdots & b_{nn} \end{bmatrix}$$

式中

$$\begin{cases} b_{ii} = l_{ii}^{-1}, & i = 1, 2, \dots, n \\ b_{ij} = -l_{jj}^{-1} \sum_{k=j+1}^i b_{ik}l_{kj} \\ j = i-1, i-2, \dots, 1; i = 2, 3, \dots, n \end{cases}$$

7.4 在标准卡尔曼滤波过程中,

$$\mathbf{P}_{k,k} \approx \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

是一个奇异矩阵,但是其“平方根”

$$\mathbf{P}_{k,k}^{1/2} = \begin{bmatrix} \epsilon/\sqrt{1-\epsilon^2} & 0 \\ 0 & 1 \end{bmatrix} \approx \begin{bmatrix} \epsilon & 0 \\ 0 & 1 \end{bmatrix}$$

是一个非奇异矩阵。

7.5 类似于练习 7.1, 设 $\mathbf{A} = [a_{ij}]_{n \times n}$ 和 $\mathbf{A}^u = [l_{ij}]_{n \times n}$, 由 $\mathbf{A} = \mathbf{A}^u (\mathbf{A}^u)^T$, 可得

$$a_{ii} = \sum_{k=i}^n l_{ik}^2, \quad i = 1, 2, \dots, n$$

$$a_{ij} = \sum_{k=j}^n l_{ik}l_{jk}, \quad j \neq i; \quad i, j = 1, 2, \dots, n$$

因此可以证明:

$$l_{ii} = (a_{ii} - \sum_{k=i+1}^n l_{ik}^2)^{1/2}, \quad i = 1, 2, \dots, n$$

$$l_{ij} = (a_{ij} - \sum_{k=j+1}^n l_{ik}l_{jk})/l_{jj}, \quad j = i+1, \dots, n; \quad i = 1, 2, \dots, n$$

$$l_{ij} = 0, \quad j = 1, 2, \dots, i-1; \quad i = 2, 3, \dots, n$$

这就给出了上三角矩阵 \mathbf{A}^u 。

7.6 新公式和我们在这一章学习的一样,只是每个有上标 c 的下三角矩阵必须用相应的上标为 u 的上三角矩阵代替。

7.7 新公式和 7.3 节给出的一样,只是每个有上标 c 的下三角矩阵必须用相应的上标为 u 的上三角矩阵代替。

第 8 章

8.1 (a) 因为 $r^2 = x^2 + y^2$, 有 $\dot{r} = \frac{x}{r}\dot{x} + \frac{y}{r}\dot{y}$, 从而 $\dot{r} = v\sin\theta$ 和 $\ddot{r} = \dot{v}\sin\theta + v\dot{\theta}\cos\theta$ 。

另一方面,因为 $\tan\theta = y/x$,有

$$\dot{\theta} \sec^2\theta = (x\dot{y} - y\dot{x})/x^2 \text{ 或 } \dot{\theta} = \frac{x\dot{y} - y\dot{x}}{x^2 \sec^2\theta} = \frac{x\dot{y} - y\dot{x}}{r^2} = \frac{v}{r}\cos\theta, \text{使得}$$

$$\ddot{r} = a\sin\theta + \frac{v^2}{r}\cos^2\theta$$

$$\ddot{\theta} = \left(\frac{\dot{v}r - v\dot{r}}{r^2} \right) \cos\theta - \frac{v}{r}\dot{\theta}\sin\theta = \left(\frac{ar - v^2\sin\theta}{r^2} \right) \cos\theta - \frac{v^2}{r^2}\sin\theta\cos\theta$$

$$(b) \dot{\mathbf{x}} = f(\mathbf{x}) := \begin{bmatrix} v\sin\theta \\ a\sin\theta + \frac{v^2}{r}\cos^2\theta \\ (ar - v^2\sin\theta)\cos\theta/r^2 - v^2\sin\theta\cos\theta/r^2 \end{bmatrix}$$

(c)

$$\mathbf{x}_{k+1} = \begin{bmatrix} x_k[1] + hv\sin(x_k[3]) \\ x_k[2] + ha\sin(x_k[3]) + v^2\cos^2(x_k[3])/x_k[1] \\ v\cos(x_k[3])/x_k[1] \\ (ax_k[1] - v^2\sin(x_k[3]))\cos(x_k[3])/x_k[1]^2 - v^2\sin(x_k[3])\cos(x_k[3])/x_k[1]^2 \end{bmatrix} + \underline{\xi}_k$$

$$v_k = [1 \ 0 \ 0 \ 0]\mathbf{x}_k + \underline{\eta}_k$$

式中 $\mathbf{x}_k := [x_k[1] \ x_k[2] \ x_k[3] \ x_k[4]]^\top$ 。

(d) 借助式(8.8)。

8.2 证明很直接。

8.3 证明很直接,可以证出: $\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} = \mathbf{f}_{k-1}(\hat{\mathbf{x}}_{k-1})$ 。

8.4 对改进了的观测方程两边取方差

$$\mathbf{v}_0 - \mathbf{C}_0(\boldsymbol{\theta})E(\mathbf{x}_0) = \mathbf{C}_0(\boldsymbol{\theta})\mathbf{x}_0 - \mathbf{C}_0(\boldsymbol{\theta})E(\mathbf{x}_0) + \underline{\eta}_0$$

左边应用估计 $(v_0 - C_0(\theta)E(x_0))(v_0 - C_0(\theta)E(x_0))^T$ 作为方差 $\text{Var}(v_0 - C_0(\theta)E(x_0))$, 有 $(v_0 - C_0(\theta)E(x_0))(v_0 - C_0(\theta)E(x_0))^T = C_0(\theta)\text{Var}(x_0)C_0(\theta)^T + R_0$, 这就可以立即得到式(8.13)。

8.5 由于

$$E(v_1) = C_1(\theta)A_0(\theta)E(x_0)$$

对改进了的观测方程两边取方差:

$$v_1 - C_1(\theta)A_0(\theta)E(x_0) = C_1(\theta)(A_0(\theta)x_0 - C_1(\theta)A_0(\theta)E(x_0) + \Gamma(\theta)\xi_0) + \eta_1$$

左边应用估计 $(v_1 - C_1(\theta)A_0(\theta)E(x_0))(v_1 - C_1(\theta)A_0(\theta)E(x_0))^T$ 作为方差 $\text{Var}(v_1 - C_1(\theta)A_0(\theta)E(x_0))$, 有

$$(v_1 - C_1(\theta)A_0(\theta)E(x_0))(v_1 - C_1(\theta)A_0(\theta)E(x_0))^T$$

$$= C_1(\theta)A_0(\theta)\text{Var}(x_0)A_0^T(\theta)C_1^T(\theta) + C_1(\theta)\Gamma_0(\theta)Q_0\Gamma_0^T C_1^T(\theta) + R_1$$

这就可以得到式(8.14)。

8.6 直接应用式(8.8)。

8.7 因为 $\underline{\theta}$ 是常值向量, 有 $S_k := \text{Var}(\underline{\theta}) = \mathbf{0}$, 所以

$$P_{0,0} = \text{Var}\begin{pmatrix} x \\ \underline{\theta} \end{pmatrix} = \begin{bmatrix} \text{Var}(x_0) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

根据简单的代数运算, 有

$$P_{k,k-1} = \begin{bmatrix} * & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad G_k = \begin{bmatrix} * \\ \mathbf{0} \end{bmatrix}$$

其中*表示矩阵中常值分块矩阵。

则式(8.15)的最后公式给出 $\hat{\theta}_{k|k} \equiv \hat{\theta}_{k-1|k-1}$ 。

8.8

$$\left\{ \begin{array}{l} \begin{aligned} \begin{bmatrix} \hat{x}_0 \\ \hat{c}_0 \end{bmatrix} &= \begin{bmatrix} x^0 \\ c^0 \end{bmatrix} \\ P_{0,0} &= \begin{bmatrix} p_0 & 0 \\ 0 & s_0 \end{bmatrix} \\ P_{k,k-1} &= P_{k-1,k-1} + \begin{bmatrix} q_{k-1} & 0 \\ 0 & s_{k-1} \end{bmatrix} \\ G_k &= P_{k,k-1} \begin{bmatrix} \hat{c}_{k-1} \\ 0 \end{bmatrix} \left[\begin{bmatrix} \hat{c}_{k-1} & 0 \end{bmatrix} P_{k,k-1} \begin{bmatrix} \hat{c}_{k-1} \\ 0 \end{bmatrix} + r_k \right]^{-1} \\ P_{k,k} &= [I - G_k \begin{bmatrix} \hat{c}_{k-1} & 0 \end{bmatrix}] P_{k,k-1} \\ \begin{bmatrix} \hat{x}_k \\ \hat{c}_k \end{bmatrix} &= \begin{bmatrix} \hat{x}_{k-1} \\ \hat{c}_{k-1} \end{bmatrix} + G_k (v_k - \hat{c}_{k-1} \hat{x}_{k-1}) \\ k &= 1, 2, \dots \end{array} \end{array} \right.$$

其中 c^0 是 \hat{c}_0 根据式(8.13)得到的估计; 即满足

$$v_0^2 - 2v_0x^0c^0 + [(x^0)^2 - p_0](c^0)^2 - r_0 = 0$$

第9章

9.1 (a) 设 $\vec{x}_k = [x_k \quad \dot{x}_k]^T$, 则

$$\begin{cases} x_k = -(\alpha + \beta - 2)x_{k-1} - (1 - \alpha)x_{k-2} + \alpha v_k + (-\alpha + \beta)v_{k-1} \\ \dot{x}_k = -(\alpha + \beta - 2)\dot{x}_{k-1} - (1 - \alpha)\dot{x}_{k-2} + \frac{\beta}{h}v_k - \frac{\beta}{h}v_{k-1} \end{cases}$$

$$(b) 0 < \alpha < 1 \text{ 和 } 0 < \beta < \frac{\alpha^2}{1-\alpha}.$$

9.2 系统(9.11)的解耦公式直接通过代数运算得到。

9.3 (a)

$$\Phi = \begin{bmatrix} 1 - \alpha & (1 - \alpha)h & (1 - \alpha)h^2/2 & -s\alpha \\ -\beta/h & 1 - \beta & h - \beta h/2 & -s\beta/h \\ -\gamma/h^2 & 1 - \gamma/h & 1 - \gamma/2 & -s\gamma/h^2 \\ -\theta & -\theta/h & -\theta h^2/2 & s(1 - \theta) \end{bmatrix}$$

(b)

$$\begin{aligned} \det[zI - \Phi] &= z^4 + [(\alpha - 3) + \beta + \gamma/2 - (\theta - 1)s]z^3 \\ &\quad + [(3 - 2\alpha) - \beta + \gamma/2 + (3 - \alpha - \beta - \gamma/2 - 3\theta)s]z^2 \\ &\quad + [(\alpha - 1) - (3 - 2\alpha - \beta + \gamma/2 - 3\theta)s]z + (1 - \alpha - \theta)s \end{aligned}$$

$$\tilde{\mathbf{X}}_1 = \frac{z\mathbf{V}(z-s)}{\det[zI - \Phi]} \{ \alpha z^2 + (\gamma/2 + \beta - 2\alpha)z + (\gamma/2 - \beta + \alpha) \}$$

$$\tilde{\mathbf{X}}_2 = \frac{z\mathbf{V}(z-1)(z-s)}{\det[zI - \Phi]} \{ \beta z - \beta + \gamma \} / h$$

$$\tilde{\mathbf{X}}_3 = \frac{z\mathbf{V}(z-1)^2(z-s)}{\det[zI - \Phi]} \gamma / h^2$$

$$\mathbf{W} = \frac{z\mathbf{V}(z-1)^3}{\det[zI - \Phi]} \theta$$

(c) 设 $\check{\mathbf{X}}_k = [x_k \quad \dot{x}_k \quad \ddot{x}_k \quad w_k]^T$, 则

$$\begin{aligned} x_k &= a_1 x_{k-1} + a_2 x_{k-2} + a_3 x_{k-3} + a_4 x_{k-4} + \alpha v_k \\ &\quad + (-2\alpha - s\alpha + \beta + \gamma/2)v_{k-1} \\ &\quad + [\alpha - \beta + \gamma/2 + (2\alpha - \beta - \gamma/2)s]v_{k-2} - (\alpha - \beta + \gamma/2)s v_{k-3} \end{aligned}$$

$$\begin{aligned} \dot{x}_k &= a_1 \dot{x}_{k-1} + a_2 \dot{x}_{k-2} + a_3 \dot{x}_{k-3} + a_4 \dot{x}_{k-4} + (\beta/h)v_k \\ &\quad - [(2+s)\beta/h - (\gamma/h)]v_{k-1} + [\beta/h - \gamma/h + (2\beta - \gamma)s/h]v_{k-2} \\ &\quad - [(\beta - \gamma)s/h]v_{k-3} \end{aligned}$$

$$\begin{aligned}\ddot{x}_k &= a_1 \ddot{x}_{k-1} + a_2 \ddot{x}_{k-2} + a_3 \ddot{x}_{k-3} + a_4 \ddot{x}_{k-4} + (\gamma/h)v_k \\ &\quad - [(2+\gamma)\gamma/h^2]v_{k-1} + (1+2s)v_{k-2} - sv_{k-3} \\ w_k &= a_1 w_{k-1} + a_2 w_{k-2} + a_3 w_{k-3} + a_4 w_{k-4} \\ &\quad + (\gamma/h^2)(v_k - 3v_{k-1} + 3v_{k-2} - v_{k-3})\end{aligned}$$

满足初始条件 $x_{-1} = \dot{x}_{-1} = \ddot{x}_{-1} = w_0 = 0$, 其中

$$\begin{aligned}a_1 &= -\alpha - \beta - \gamma/2 + (\theta - 1)s + 3 \\ a_2 &= 2\alpha + \beta - \gamma/2 + (\alpha + \beta h + \gamma/2 + 3\theta - 3)s - 3 \\ a_3 &= -\alpha + (-2\alpha - \beta + \gamma/2 - 3\theta + 3)s + 1 \\ a_4 &= (\alpha + \theta - 1)s\end{aligned}$$

(d) 可以直接证明。

9.4 证明很繁琐但是很简单。

9.5 研究式(9.19)和式(9.20)。必须有 $\sigma_p, \sigma_v, \sigma_a \geq 0, \sigma_m > 0$ 和 $P > 0$ 。

9.6 该方程可以通过简单的代数运算得到。

9.7 只需要简单的代数运算即可得到。

第10章

10.1 对于(1)~(4), 设 $* \in \{+, -, \cdot, /\}$, 则

$$\begin{aligned}X * Y &= \{x * y \mid x \in X, y \in Y\} \\ &= \{y * x \mid y \in Y, x \in X\} \\ &= Y * X\end{aligned}$$

其他的可以通过类似的方式证明。至于(7)的(c)部分, 不失一般性, 可以只考虑 $X = [\underline{x}, \bar{x}]$ 和 $Y = [\underline{y}, \bar{y}]$ 中 $x \geq 0$ 和 $y \geq 0$ 的情形, 然后讨论 $\underline{z} \geq 0, \bar{z} \leq 0$ 和 $\underline{z} \bar{z} < 0$ 等不同情况。

10.2 通过定义可以直接证明所有的公式。例如, 对于(j.1), 有

$$\begin{aligned}\mathbf{A}^T(\mathbf{BC}) &= \left[\sum_{j=1}^n \mathbf{A}^T(i, j) \left[\sum_{l=1}^n \mathbf{B}_{jl} \mathbf{C}_{lk} \right] \right] \\ &\subseteq \left[\sum_{j=1}^n \sum_{l=1}^n \mathbf{A}^T(i, j) \mathbf{B}_{jl} \mathbf{C}_{lk} \right] \\ &= \left[\sum_{l=1}^n \left[\sum_{j=1}^n \mathbf{A}^T(i, j) \mathbf{B}_{jl} \right] \mathbf{C}_{lk} \right] \\ &= (\mathbf{A}^T \mathbf{B}) \mathbf{C}\end{aligned}$$

10.3 见 Alefeld, G. 和 Herzberger, J. (1983)。

10.4 与练习 1.10 类似。

10.5 注意到有界系统和其任意邻近系统滤波结果曲线会不时互相交错。

10.6 见 Siouris, G., Chen, G. 和 Wang, J. (1997)。

第 11 章

11.1

$$\phi_2(t) = \begin{cases} \frac{1}{2}t^2, & 0 \leq t < 1 \\ -t^2 + 3t - \frac{3}{2}, & 1 \leq t < 2 \\ \frac{1}{2}t^2 - 3t + \frac{9}{2}, & 2 \leq t < 3 \\ 0, & \text{其他} \end{cases}$$

$$\phi_3(t) = \begin{cases} \frac{1}{6}t^3, & 0 \leq t < 1 \\ -\frac{1}{2}t^3 + 2t^2 - 2t + \frac{2}{3}, & 1 \leq t < 2 \\ \frac{1}{2}t^3 - 4t^2 + 10t - \frac{22}{3}, & 2 \leq t < 3 \\ -\frac{1}{6}t^3 + 2t^2 - 8t + \frac{32}{3}, & 3 \leq t < 4 \\ 0, & \text{其他} \end{cases}$$

11.2

$$\hat{\phi}_n(\omega) = \left(\frac{1 - e^{-i\omega}}{i\omega} \right)^n = e^{-in\omega/2} \left(\frac{\sin(\omega/2)}{\omega/2} \right)^n$$

11.3 简单的示意图。

11.4 直接代数运算。

11.5 直接代数运算。

索引

| | | |
|-------------------------|---|------------|
| ARMA(自回归滑动平均)过程 | ARMA(autoregressive moving-average) process | 25 |
| Cholesky 分解 | Cholesky factorization | 183 |
| FIR 系统 | FIR system | 158 |
| IIR 系统 | IIR system | 158 |
| LU 分解 | LU decomposition | 161 |
| Schur 互补技术 | Schur complement technique | 161 |
| z -变换 | z -transform | 114 |
| z -反变换 | inverse z -transform | 115 |
| $\alpha\beta\gamma$ 跟踪器 | $\alpha\beta\gamma$ tracker | 117 |
| $\alpha\beta$ 跟踪器 | $\alpha\beta$ tracker | 121 |
| 白噪声序列(过程) | white noise sequence (process) | 12 |
| 贝叶斯公式 | Bayes formula | 10 |
| 边缘概率密度函数 | marginal probability density function | 9,10 |
| 并行处理 | parallel processing | 161 |
| 参数辨识 | parameter identification | 93 |
| “次优”滤波器 | suboptimal filter | 117 |
| 笛卡儿正负号法则 | Descartes rule of signs | 121 |
| 方差 | variance | 8,43,98 |
| 方位角误差 | azimuthal angular error | 39 |
| 仿射模型 | affine model | 42-44 |
| 非负定矩阵 | non-negative definite matrix | 23,42,47 |
| 非线性模型(系统) | nonlinear model (system) | 93,96,98 |
| 分离原理 | separation principle | 160 |
| 改进扩展卡尔曼滤波 | modified extended Kalman filter | 93 |
| 概率论初步 | probabliby preliminaries | 6 |
| 概率分布 | probability distribution | 7,8 |
| 概率密度函数 | probabliby density function | 7,8 |
| 高斯(或正态)概率密度函数 | Gaussian (or normal) probability density function | 9 |
| 高斯(或标准)白噪声序列 | Gaussian (or normal) white noise sequence | 12 |
| 高斯白噪声序列 | Gaussian white noise sequence | 98,122 |
| 估计 | estimate | 13,103,180 |
| 函数 | function | 6,129,159 |
| 迹 | trace | 4,167 |

| | | |
|----------------|--|-------------|
| 基本事件 | simple event | 6 |
| 极限(或稳态)卡尔曼滤波器 | limiting (or steady-state) Kalman filter | 113 |
| 极限卡尔曼滤波器 | limiting Kalman filter | 118,120,121 |
| 极限卡尔曼增益矩阵 | limiting Kalman gain matrix | 69 |
| 几何收敛 | geometric convergence | 77 |
| 渐近最优估计 | asymptotically optimal estimate | 78 |
| 角位移 | angular displacement | 96 |
| 解耦公式 | decoupling formulas | 113,121 |
| 矩阵 Riccati 方程 | matrix Riccati equation | 82 |
| 矩阵求逆引理 | matrix inversion lemma | 2 |
| 矩阵施瓦兹不等式 | matrix Schwarz inequality | 2 |
| 距离、范围 | range | 124,125,160 |
| 卡尔曼-布希滤波器 | Kalman-Bucy filter | 158 |
| 卡尔曼滤波方程(算法,过程) | Kalman filtering equation(algorithm, or process) | 31,47 |
| 卡尔曼滤波器 | Kalman filter | 1,141 |
| 卡尔曼平滑器 | Kalman smoother | 152 |
| 卡尔曼增益矩阵 | Kalman gain matrix | 19,21,48 |
| 克莱姆法则 | Cramer's rule | 116,122 |
| 扩展卡尔曼滤波器 | extended Kalman filter | 93 |
| 雷达跟踪模型(系统) | radar tracking model(or system) | 39 |
| 联合概率分布(函数) | joint probability distribution (function) | 9 |
| 联合概率密度函数 | joint probability density function | 9 |
| 零均值高斯白噪声序列 | zero-mean Gaussian white noise sequence | 17,23 |
| 滤波方程解耦 | decoupling of filtering equation | 113 |
| 平方根矩阵 | square-root matrix | 162 |
| 平方根算法 | square-root algorithm | 84,88 |
| 期望 | expectation | 8,13,130 |
| 区间卡尔曼滤波 | interval Kalman filter | 132 |
| 权值矩阵 | weight matrix | 13 |
| 确定性输入序列 | deterministic input sequence | 16 |
| 施瓦兹不等式 | Schwarz inequality | 1,2 |
| 实时跟踪 | real-time tracking | 1,36 |
| 实时应用 | real-time application | 52 |
| 实时应用算法 | algorithm for real-time application | 91 |
| 实现 | implementation | 17,95,160 |
| 事件 | event | 6 |
| 输出 | outcome | 6 |
| 数字滤波过程 | digital filtering process | 18 |
| 数字平滑过程 | digital smoothing process | 18 |

| | | |
|----------------------------|--|-------------|
| 数字预测过程 | digital prediction process | 18 |
| 随机变量 | random variable | 7 |
| 随机向量 | random vector | 8,9 |
| 随机信号 | random signal | 141,146 |
| 随机序列 | random sequence | 12 |
| 随机最优控制 | stochastic optimal control | 159 |
| 泰勒近似 | Taylor approximation | 93 |
| 条件方差 | conditional variance | 11,132 |
| 条件概率 | conditional probability | 9,10 |
| 条件概率密度函数 | conditional probability density function | 10 |
| (条件)期望 | conditional expectation | 133 |
| 外源自回归滑动平均模型(带输入的自回归滑动平均模型) | ARMAX (autoregressive moving-average model with exogeneous inputs) | 57 |
| 维纳滤波器 | Wiener filter | 158 |
| 卫星轨道估计 | satellite orbit estimation | 96 |
| 极限卡尔曼滤波器 | limiting Kalman filter | 118,120,121 |
| 稳态卡尔曼滤波器 | steady-state Kalman filter | 68,69 |
| 无偏估计 | unbiased estimate | 14,43,46 |
| 吸引点 | attracting point | 96 |
| 系统噪声和量测噪声过程 | correlated system and measurement noise processes | 42 |
| 线性调节器问题 | linear regulator problem | 159 |
| 线性估计 | linear estimate | 14 |
| 线性确定/随机系统 | linear deterministic/stochastic system | 39,54 |
| 向量形式的施瓦兹不等式 | vector Schwarz inequality | 2 |
| 小波 | wavelets | 141,143 |
| 小波卡尔曼滤波器 | wavelet Kalman filter | 141 |
| 协方差 | covariance | 10,11 |
| 新息序列 | innovations sequence | 29 |
| 行列式初步 | determinant preliminaries | 1 |
| 序贯算法 | sequential algorithm | 84 |
| 仰角误差 | elevational angular error | 39 |
| 有色噪声(序列、过程) | colored noise (sequence or process) | 58,63 |
| 预测-校正 | prediction-correction | 16,18 |
| 若尔当标准型 | Jordan canonical form | 4 |
| 噪声自适应滤波 | noise-adaptive filter | 157 |
| 增广矩阵 | augmented matrix | 161 |
| 增广系统 | augmented system | 67 |
| 正定矩阵 | positive definite matrix | 9,13 |
| 正态分布 | normal distribution | 7 |

| | | |
|----------|---------------------------------|---------|
| 自适应卡尔曼滤波 | adaptive Kalman filtering | 156,157 |
| 自适应系统辨识 | adaptive system identification | 93,97 |
| 最小二乘初步 | least-squares preliminaries | 12 |
| 最小二乘最优估计 | least-squares optimal estimate | 14 |
| 最小方差估计 | minimum variance estimate | 14,30 |
| 最小迹方差估计 | minimum trace variance estimate | 44 |
| 最优估计 | optimal estimate | 13,27 |
| 最优估计算子 | optimal estimate operator | 44 |
| 最优权值矩阵 | optimal weight matrix | 13 |
| 最优预测 | optimal prediction | 18,153 |
| 最优准则 | optimality criterion | 17 |

清华大学出版社数字出版网站



www.wqbook.com

ISBN 978-7-302-30907-9

A standard barcode for the ISBN 978-7-302-30907-9.

9 787302 309079 >

定价：32.00元

[General Information]

书名=????????????? ?4?

页数=193

SS号=13260033