先进滤波方法及其在导航中的应用

Advanced Filtering Methods and Their Applications in Navigation

宁晓琳　宫晓琳　李建利　编著

国防工业出版社
National Defense Industry Press

先进滤波方法及其在导航中的应用

宁晓琳　宫晓琳　李建利　编著

国防工業出版社
·北京·

内容简介

本书是作者课题组多年来的教学和科研成果的总结，主要介绍了线性系统的卡尔曼滤波和非线性系统的扩展卡尔曼滤波、Unscented 卡尔曼滤波等滤波方法和这些滤波方法在 INS 初始对准、GPS 动态滤波、GPS/INS 组合导航以及 GPS/INS/CNS 组合导航中的应用。

本书既可作为从事导航与制导技术研究的工程技术人员的参考书，也可作为高等学校相关专业研究生和高年级本科生的教材或教学参考书。

图书在版编目（CIP）数据

先进滤波方法及其在导航中的应用/宁晓琳，宫晓琳，李建利编著. —北京：国防工业出版社，2019. 4

ISBN 978-7-118-11659-5

Ⅰ. ①先… Ⅱ. ①宁… ②宫… ③李… Ⅲ. ①滤波技术—应用—卫星导航 Ⅳ. ①TN967. 1

中国版本图书馆 CIP 数据核字（2019）第 038295 号

※

国防工业出版社出版发行

（北京市海淀区紫竹院南路 23 号 邮政编码 100048）

三河市众誉天成印务有限公司印刷

新华书店经售

*

开本 710×1000 1/16 **印张** 9½ **字数** 175 千字

2019 年 4 月第 1 版第 1 次印刷 **印数** 1—2000 册 **定价** 78.00 元

国防书店：（010）88540777　　发行邮购：（010）88540776

发行传真：（010）88540755　　发行业务：（010）88540717

序

导航技术是指利用各类导航传感器获得的测量信息得到运载体的位置、速度和姿态等运动参数。运行于海陆空天上的各类运载体，包括舰船、汽车、飞机和卫星等，其正常工作都离不开导航技术。导航技术根据其测量信息的不同，可分为惯性导航、卫星导航、天文导航、视觉导航、地球物理场导航等。精度是衡量导航系统性能的重要指标，在敏感器精度无法提高的条件下，如何有效地抑制和降低测量误差、器件误差和模型误差等各类误差或利用不同导航方式之间的互补性进行组合导航就成为提高导航精度的主要手段。而误差抑制和组合导航都需要用到卡尔曼滤波等先进的滤波方法，因此先进滤波和组合导航是导航技术的核心内容。

针对我国对导航领域人才的迫切需求，我于20世纪90年代在北京航空航天大学仪器学科开设了研究生专业基础课“卡尔曼滤波与组合导航”，该课程主要介绍卡尔曼滤波、Unscented 卡尔曼滤波、粒子滤波等先进滤波方法的基本原理及其在惯性导航、卫星导航、天文导航和组合导航中的应用和该领域国内外最新的研究进展，为研究现代导航理论和导航技术打下基础。课程开设后受到了学生们的欢迎，目前已经成为北京航空航天大学的研究生精品课程。

课程开设之初，采用由秦永元、张洪钺、汪叔华三位先生编著的《卡尔曼滤波与组合导航原理》作为主要参考书，该教材内容循序渐进，由浅入深，对于滤波方法的公式和定理附有详细推导和证明，是一本不可多得的经典教材。但是近年来，随着我国航天和国防对高性能导航技术的迫切需求，各种导航技术飞速发展，基于新概念、新原理、新传感器和新应用的导航技术不断涌现。例如，在惯性导航方向，高精度的光学陀螺和低成本的MEMS陀螺的获得了快速发展和应用；在卫星导航方面，我国拥有完全自主知识产权、自主建设、独立运行并与世界其他卫星导航系统相兼容的北斗卫星导航系统部署成功并广泛应用于国民经济的方方面面；在天文导航方面，我国月球探测、火星探测和其他深空探测任务相继实施和开展，为天文导航的发展提供了新的应用领域和发展方向。另外，各种无人机、无人车、无人

潜器、机器人等无人系统的不断发展也对导航系统和导航方法提出了更高的要求。为了适应导航技术日新月益的发展，教学团队适时调整教学内容，及时地汲取国内外导航技术的最新进展，并成功地将团队的科研成果转化为教学内容，以符合不断发展的教学需要。为了配合教学要求，增加了团队的系列专著，包括《惯性导航初始对准》《GPS 动态滤波的理论、方法及其应用》《惯性/天文/卫星组合导航技术》《航天器天文导航原理与方法》《深空探测器自主天文导航方法》等作为教学参考书，但导致参考书目较多，选修本门课程学生迫切需要一本适合本课程的教材的。为此，科研团队三位年轻教师宁晓琳、宫晓琳和李建利，将团队在多年的教学过程中通过不断积累，结合教学实际和科研经验形成的课程讲义重新进行了梳理，编写完成了本教材。本教材的特点在于基础性、实用性和先进性的统一，既包含滤波方法和组合导航相关基础知识、基本理论和基本方法，又密切联系工程实际，通过导航实例，方便学生学以致用，同时吸纳了该领域的最新科研成果，体现了国内外最新的研究进展。

希望这本教材能够作为《卡尔曼滤波与组合导航原理》课程的辅助教材，同时为我国导航领域的科研人员提供参考，并为我国导航技术的研究和应用起到积极的促进作用。

房建成

2019.1.27

前言

导航是指利用各类测量信息获得载体的位置、速度和姿态等参数的技术，导航技术是舰艇、飞机和航天器等各类运载体必不可少的关键技术，在国防和民用领域均具有重要的作用。无论是单一的导航系统还是组合导航系统都会受到各种误差的影响，为了抑制误差，提高导航精度，就需要用到各种先进的滤波方法，因此滤波方法在导航中有广泛的应用，例如惯性导航的初始对准、卫星导航的信息动态滤波和航天器的天文导航等。而组合导航的信息分配和信息融合也主要是利用滤波方法实现的，因此各类组合导航系统如惯性/卫星、惯性/天文、惯性/地磁、惯性/视觉等均离不开各种滤波方法。

本教材来源于作者承担的北京航空航天大学研究生课程《卡尔曼滤波与组合导航》的课程讲义，在编写时参考了秦永元、张洪钺、汪叔华三位先生编著的《卡尔曼滤波与组合导航原理》及国内多本介绍先进滤波方法的教材和专著，同时融入了作者多年的教学经验和团队的部分科研成果。主要包括5章，其中第1章首先对滤波的作用、卡尔曼滤波的发展历史、卡尔曼滤波理论在导航中的应用和状态模型和量测模型建立方法进行了介绍。第2章系统地介绍了最小二乘估计、最小方差估计、线性离散系统卡尔曼滤波和连续系统卡尔曼滤波等显性系统的卡尔曼滤波方法。第3章介绍了针对非线性系统的滤波方法、Sigma点卡尔曼滤波方法和粒子滤波方法。在此基础上，第4章介绍了滤波方法在惯性导航系统（Inertial Navigation System，INS）、卫星导航（Global Positioning System，GPS）和天文导航（Celestial Navigation System，CNS）中的应用。第5章则介绍了滤波方法在GPS/INS组合导航以及GPS/INS/CNS组合导航中的应用。本教材可作为高等学校相关专业研究生和高年级本科生的教材或教学参考书，也可供从事导航技术研究的工程技术人员的参考。

本书第1~第3章和4.3节由宁晓琳撰写，第5章和4.2节由宫晓琳撰写，第4.1节由李建利撰写。由于作者水平和时间有限，难免存在不妥和错误之处，恳请广大同行、读者批评指正。

作者

2019年1月

目录

第1章 绪论

1.1 滤波的作用

滤波一词起源于通信理论,其本质是从含有误差(噪声、干扰)的信号中提取有用信号的一种技术。像传统的高通、低通、带通、带阻滤波器一样,卡尔曼滤波(KF)也可以实现对误差的抑制,但不同的是传统滤波器只能在信号与误差具有不同频带的条件下才能实现,而卡尔曼滤波则没有这个限制,但卡尔曼滤波也只能处理随机误差。

虽然误差的来源多种多样,但从性质上主要可分为系统误差和随机误差两类。系统误差又称为可测误差或规律误差,它是由于实验方法、所用仪器、测量原理等因素造成的按某些确定规律变化的误差。这类误差的特征是数值保持恒定,或遵循一定的规律变化。系统误差可通过建模标定的方法进行补偿。随机误差又称偶然误差,是指其误差的数值不固定,没有规律可循。随机信号是不能用确定的数学关系式来描述的,不能预测其未来任何瞬时值。虽然单次测量的随机误差没有规律,但多次测量的总体却服从统计规律,因此可利用卡尔曼滤波等最优估计方法进行抑制。

1.2 卡尔曼滤波的发展历史

1801 年,高斯(Johann Carl Friedrich Gauss)为了测定小行星谷神星(Ceres)的轨道提出了最小二乘估计法(LS),该方法不需要任何先验知识,但对时变系统的估计精度相对较低。在第二次世界大战期间,维纳(Norbert Wiener) 为了解决防空火力控制和雷达噪声滤波问题,于 1942 年 2 月首先给出了从时间序列的过去测量数据推知未来信号的维纳滤波,建立了在最小均方误差准则下将时间序列外推进行预测的维纳滤波理论[1]。柯尔莫哥罗夫(Andrey Kolmogorov)于 1941 年独立

推导出了与维纳滤波等效的离散系统的滤波方法,因此维纳滤波方法也称为维纳-柯尔莫哥罗夫滤波。维纳滤波器的优点是适应面较广,无论平稳随机过程是连续的还是离散的都可应用,并且对某些问题还可求出滤波器传递函数的显式解,并进而采用由简单的物理元件组成的网络构成维纳滤波器。其缺点是要求信号和噪声均是平稳的线性随机过程,只能在频域使用,并需要利用全部测量数据求解,计算量大,因此维纳滤波在实际问题中应用不多。卡尔曼滤波(Kalman Filter, KF)以它的发明者卡尔曼(Rudolph E. Kalman)命名,但实际上蒂勒(Thorvald Nicolai Thiele)和斯沃林(Peter Swerling)在更早之前就提出了类似的算法[2,3]。由于南加州大学的布西(Richard S. Bucy)对该方法也做出了很大贡献,因此该滤波方法也常被称为卡尔曼-布西滤波。卡尔曼在美国国家航空航天局(NASA)埃姆斯研究中心访问时,施密特(Stanley F. Schmidt)发现卡尔曼滤波对于解决“阿波罗”计划中的非线性轨道预测很有用,因此在“阿波罗”飞船的导航程序便使用了这种滤波器,这也是卡尔曼滤波的第一个实际应用[4]。卡尔曼滤波不要求信号和噪声都是平稳过程的假设条件,对于每个时刻的系统扰动和观测误差,只要对它们的统计性质作某些适当的假设,通过对含有噪声的观测信号进行处理,就能在平均的意义上,求得误差为最小的真实信号的估计值。

由于卡尔曼滤波要求系统为线性系统,噪声为高斯噪声且统计特性明确可知,并且在计算中涉及矩阵求逆,因此在实际应用中出现了很多扩展方法。如解决非线性系统估计问题的扩展卡尔曼滤波(Extended Kalman Filter, EKF)、无迹卡尔曼滤波(Unscented Kalman Filter, UKF)、中心差分卡尔曼滤波(Central Difference Kalman Filter, CDKF)和容积卡尔曼滤波(Cubature Kalman Filter, CKF)。解决估计误差方差阵不正定问题的平方根滤波、UDU^T分解滤波、信息滤波。解决噪声统计特性时变或不准确的自适应滤波、多模型滤波,解决非高斯噪声的粒子滤波(Particle Filter, PF)等。

1.3 卡尔曼滤波理论在导航中的应用

导航是指利用各类测量信息获得载体的位置、速度和姿态等参数的技术。导航通常与制导和控制一起构成一个闭环系统,共同完成使载体从起始点到达目标点的任务。根据所使用量测信息的不同,导航系统可分为惯性导航系统、卫星导航系统、天文导航系统、物理场(磁场、重力场)匹配导航系统、视觉(图像)导航系统等。各种导航系统各有自己的特点,例如,惯性导航具有自主性强、短时间精度高、可连续提供位置、速度、姿态的优点,但缺点是误差随时间积累,且精度越高价格越昂贵。卫星导航的优点是精度高,误差不积累,可全球、全天时、全天候提供导航信息,并且接收机价格便宜,但缺点是易受干扰和破坏,不能输出姿态信息,且输出不

连续。天文导航完全自主,误差不积累,并且可同时提供位置和姿态信息,但缺点是导航精度受天体敏感器精度制约,总体定位精度不够高,且输出信息不连续。由于每种导航方法都有自己的优势,也有不足,因此为了获得更好的导航性能,常通过将几种不同的导航方法进行组合构成组合导航系统,从而达到优势互补的目的。

无论是单一的导航系统还是组合导航系统都会受到各种误差的影响,为了抑制误差,提高导航精度,就需要用到卡尔曼滤波,因此卡尔曼滤波在导航中有广泛的应用。例如,惯性导航的初始对准,卫星导航的信息动态滤波,航天器的天文导航。而组合导航的信息分配和信息融合更主要是利用卡尔曼滤波实现的,因此各类组合导航系统如惯性/卫星、惯性/天文、惯性/地磁、惯性/视觉等均离不开卡尔曼滤波。

从载体上说,无论是海里的舰船、地面行驶的车辆、空中飞行的飞机、导弹还是太空中的航天器,其导航也离不开卡尔曼滤波。例如,美国海军的“战斧”导弹、美国空军的空基巡航导弹、可重复使用运载器、国际空间站的导航系统中均使用了卡尔曼滤波,可以说卡尔曼滤波是目前各类载体导航系统中最重要的最优估计方法。

1.4 状态模型和量测模型

使用卡尔曼滤波的前提是要建立两个模型:状态模型和量测模型。状态量通常就是要被估计的量。在导航系统中,一般将位置、速度和姿态或它们的误差作为状态量。状态模型是描述状态量随时间变化规律的数学模型。量测量则是指由敏感器输出的量测信息,例如,惯性敏感器输出的加速度和角速度,卫星导航接收机输出的位置、速度、伪距、伪距率,天体敏感器输出的天体高度和方向等。量测模型则是反映状态量和量测量之间关系的数学模型。

如果状态模型没有误差,则在初始位置、速度和姿态已知的情况下,就可以根据状态模型预测得到任意时刻的位置、速度和姿态,完全不需量测量的帮助。如果量测量完全没有误差,则也可以根据当前时刻的量测量和量测模型反推出当前时刻的位置、速度和姿态,也不需要状态模型的帮助。但是,在实际应用中,无论是状态模型还是量测量都不可避免地存在误差。例如,卫星的轨道可以用轨道动力学模型描述,也可通过地面站对其进行实时观测,但无论轨道动力学还是地面站的量测都不可能没有误差。这时就需要利用卡尔曼滤波对状态模型和量测量的误差进行抑制,获得载体导航信息的最优估计。

设被估计状态为 $\boldsymbol{X}$, 状态模型噪声为 $\boldsymbol{W}$, 量测量为 $\boldsymbol{Z}$, 量测噪声为 $\boldsymbol{V}$, 则状态方程和量测方程的连续形式和离散形式通常可表示如下:

$$\begin{cases}\dot{\boldsymbol{X}}(t)=\varphi[\boldsymbol{X}(t),t]+\boldsymbol{\Gamma}[\boldsymbol{X}(t),t]\boldsymbol{W}(t)\\ \boldsymbol{Z}(t)=h[\boldsymbol{X}(t),t]+\boldsymbol{V}(t)\end{cases}\tag{1-1}$$

$$
\begin{cases} \boldsymbol{X}(k+1)=\varphi[\boldsymbol{X}(k),k]+\boldsymbol{\Gamma}[\boldsymbol{X}(k),k]\boldsymbol{W}(k) \\ \boldsymbol{Z}(k+1)=h[\boldsymbol{X}(k+1),k+1]+\boldsymbol{V}(k+1) \end{cases} \tag{1-2}
$$

式中：$\varphi(\cdot)$，$h(\cdot)$为已知的向量函数；$\boldsymbol{\Gamma}(\cdot)$为系统噪声驱动矩阵。

下面通过几个例子，介绍如何建立系统的状态模型和量测模型。

例 1 一个小车在地面静止不动，车上装有一个可提供位置的 GPS 接收机，如果想利用卡尔曼滤波技术估计小车的位置，则该怎样建立其状态模型和量测模型？

解：令状态量 $\boldsymbol{X}(k)=\begin{bmatrix} x(k) \\ y(k) \\ z(k) \end{bmatrix}$，$x(k)$，$y(k)$，$z(k)$ 为小车在三维空间中的位置坐标值，由于小车静止不动，可得状态模型为

$$
\begin{bmatrix} x(k+1) \\ y(k+1) \\ z(k+1) \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x(k) \\ y(k) \\ z(k) \end{bmatrix}
$$

因为 $\boldsymbol{Z}(k)=\begin{bmatrix} x(k) \\ y(k) \\ z(k) \end{bmatrix}$，所以量测模型为

$$
\boldsymbol{Z}(k)=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\cdot\boldsymbol{X}(k)+\boldsymbol{V}=H\cdot\boldsymbol{X}(k)+\boldsymbol{V}
$$

例 2 小车做匀速直线运动，车上装有一个可提供位置的 GPS 接收机，如果利用卡尔曼滤波技术估计小车的位置，则该怎样建立其状态模型和量测模型？

解：根据匀速直线运动中，位置和速度随时间变化的规律，可建立状态模型如下，

$$
\begin{bmatrix} x(k+1) \\ y(k+1) \\ z(k+1) \\ v_x(k+1) \\ v_y(k+1) \\ v_z(k+1) \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 & T & 0 & 0 \\ 0 & 1 & 0 & 0 & T & 0 \\ 0 & 0 & 1 & 0 & 0 & T \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x(k) \\ y(k) \\ z(k) \\ v_x(k) \\ v_y(k) \\ v_z(k) \end{bmatrix}
$$

式中：$v_x(k)$，$v_y(k)$，$v_z(k)$为小车在三个方向的速度。

量测模型与例 1 类似，可得

$$
\boldsymbol{Z}(k)=H\cdot\boldsymbol{X}(k)+\boldsymbol{V}=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}\cdot\boldsymbol{X}(k)+\boldsymbol{V}
$$

例 3 小车做匀速直线运动,量测信息为某个地面站测得的到小车的相对距离,如果想利用卡尔曼滤波技术估计小车的位置,则该怎样建立其状态模型和量测模型?

解: 状态模型与例 2 相同,即

$$\begin{bmatrix} x(k+1) \\ y(k+1) \\ z(k+1) \\ v_x(k+1) \\ v_y(k+1) \\ v_z(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & T & 0 & 0 \\ 0 & 1 & 0 & 0 & T & 0 \\ 0 & 0 & 1 & 0 & 0 & T \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(k) \\ y(k) \\ z(k) \\ v_x(k) \\ v_y(k) \\ v_z(k) \end{bmatrix}$$

假设地面站的位置为 x_s、y_s、z_s,则根据距离公式,可得量测模型为

$$Z(k) = h(\boldsymbol{X}(k)) + V = \sqrt{(x(k) - x_s)^2 + (y(k) - y_s)^2 + (z(k) - z_s)^2} + V$$

例 4 小车做匀速圆周运动,量测信息为某个地面站测得的到小车的相对距离,请利用卡尔曼滤波技术估计小车的位置。

解: 根据匀速圆周运动的方程,可得状态模型为

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{v}_x \\ \dot{v}_y \\ \dot{v}_z \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \\ v_z \\ -\omega_x^2 x \\ -\omega_y^2 y \\ -\omega_z^2 z \end{bmatrix}$$

量测模型同例 3。

1.5 数学基础

本节介绍在卡尔曼滤波和其他最优估计方法的推导中常用到的公式,包括求导公式、矩阵求逆公式、泰勒公式。

1. 矩阵求导公式

(1) 设 f 为向量 $\boldsymbol{x}$ 的函数, t 为标量, $\boldsymbol{x} = [\boldsymbol{x}_1(t) \quad \boldsymbol{x}_2(t) \quad \cdots \quad \boldsymbol{x}_n(t)]^{\mathrm{T}}$,则有 $\dfrac{\mathrm{d}f}{\mathrm{d}t} = \left(\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}\right)^{\mathrm{T}} \dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$。

(2) 设f为向量$\boldsymbol{x}$的函数,$\boldsymbol{T}$为m维向量,$\boldsymbol{x}=[\boldsymbol{x}_1(\boldsymbol{T}) \quad \boldsymbol{x}_2(\boldsymbol{T}) \quad \cdots \quad \boldsymbol{x}_n(\boldsymbol{T})]^{\mathrm{T}}$
$\boldsymbol{T}=[t_1 \quad t_2 \quad \cdots \quad t_m]^{\mathrm{T}}$,则有$\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{T}}=\left[\dfrac{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}{\mathrm{d}\boldsymbol{T}}\right]\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}$。

(3) $\boldsymbol{\beta}$为n维向量,即$\boldsymbol{\beta}=[\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n]^{\mathrm{T}}$,则有$\dfrac{\mathrm{d}\boldsymbol{\beta}^{\mathrm{T}}}{\mathrm{d}\boldsymbol{\beta}}=\boldsymbol{I}_n$。

(4) $\boldsymbol{\beta}$为n维向量,$\boldsymbol{\alpha}$和$\boldsymbol{B}$为与$\boldsymbol{\beta}$无关的n维常数向量,和$n\times m$维常数矩阵,则有$\dfrac{\mathrm{d}(\boldsymbol{\alpha}+\boldsymbol{B}\boldsymbol{\beta})^{\mathrm{T}}}{\mathrm{d}\boldsymbol{\beta}}=\boldsymbol{B}^{\mathrm{T}}$。

(5) 若$\boldsymbol{A}=(\boldsymbol{\alpha}-\boldsymbol{B}\boldsymbol{\beta})^{\mathrm{T}}(\boldsymbol{\alpha}-\boldsymbol{B}\boldsymbol{\beta})$,则有$\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}\boldsymbol{\beta}}=-2\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{\alpha}-\boldsymbol{B}\boldsymbol{\beta})$。

(6) 如果$\boldsymbol{A}=\begin{bmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{bmatrix}$且$|\boldsymbol{A}|\neq 0$,则

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{A}^{-1}}{\mathrm{d}t}=-\boldsymbol{A}^{-1}\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\boldsymbol{A}^{-1} \\ \dfrac{\mathrm{d}|\boldsymbol{A}|}{\mathrm{d}t}=|\boldsymbol{A}|\operatorname{tr}\left(\boldsymbol{A}^{-1}\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\right) \end{cases}$$

(7) 设$\boldsymbol{x}$为$m\times n$维矩阵,$\boldsymbol{A}$和$\boldsymbol{B}$分别为与$\boldsymbol{x}$无关的$m\times m$和$n\times m$的常矩阵,

则$\begin{cases} \dfrac{\mathrm{d}\operatorname{tr}(\boldsymbol{B}\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}=\dfrac{\mathrm{d}\operatorname{tr}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}})}{\mathrm{d}\boldsymbol{x}}\boldsymbol{B}^{\mathrm{T}} \\ \dfrac{\mathrm{d}\operatorname{tr}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}=(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})\boldsymbol{x} \end{cases}$

(8) $\boldsymbol{x}$为n维变量,$\boldsymbol{\mu}$为n维常量,$\boldsymbol{A}$为n阶常对称矩阵,则

$$\frac{\mathrm{d}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\mu})}{\mathrm{d}\boldsymbol{x}}=2\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\mu})$$

(9) 若$\boldsymbol{B}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,则$\dfrac{\partial\boldsymbol{B}}{\partial\boldsymbol{x}}=2\boldsymbol{A}\boldsymbol{x}$。

(10) $\begin{cases} \dfrac{\partial\operatorname{tr}(\boldsymbol{B}\boldsymbol{Q})}{\partial\boldsymbol{Q}}=\boldsymbol{B}^{\mathrm{T}} \\ \dfrac{\partial\operatorname{tr}(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q})}{\partial\boldsymbol{Q}}=(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})\boldsymbol{Q} \end{cases}$

2. 矩阵求逆公式

矩阵求逆公式为

$$(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}=\boldsymbol{A}^{-1}+\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}$$

$$=-\boldsymbol{C}^{-1}\boldsymbol{D}(\boldsymbol{B}-\boldsymbol{A}\boldsymbol{C}^{-1}\boldsymbol{D})^{-1}$$

$$=-(\boldsymbol{C}-\boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{A})^{-1}\boldsymbol{D}\boldsymbol{B}^{-1}$$

3. 泰勒公式

$$(1)\ f(\boldsymbol{x})=f(\boldsymbol{x}_0)+f'(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)+\frac{f''(\boldsymbol{x}_0)}{2!}(\boldsymbol{x}-\boldsymbol{x}_0)^2+\cdots+\frac{f^{(n)}(\boldsymbol{x}_0)}{n!}(\boldsymbol{x}-\boldsymbol{x}_0)^n+\cdots+\boldsymbol{R}_n(\boldsymbol{x})$$

其中

$$\boldsymbol{R}_n(\boldsymbol{x})=\frac{f^{(n+1)}(\boldsymbol{x}_0)}{(n+1)!}(\boldsymbol{x}-\boldsymbol{x}_0)^{n+1}$$

$$(2)\ f(\boldsymbol{x}_0+\Delta\boldsymbol{x},\boldsymbol{y}_0+\Delta\boldsymbol{y})=f(\boldsymbol{x}_0,\boldsymbol{y}_0)+(\Delta\boldsymbol{x}\frac{\partial}{\partial\boldsymbol{x}}+\Delta\boldsymbol{y}\frac{\partial}{\partial\boldsymbol{y}})f'(\boldsymbol{x}_0,\boldsymbol{y}_0)+\frac{1}{2!}\left(\Delta\boldsymbol{x}\frac{\partial}{\partial\boldsymbol{x}}+\Delta\boldsymbol{y}\frac{\partial}{\partial\boldsymbol{y}}\right)^2 f(\boldsymbol{x}_0,\boldsymbol{y}_0)+\cdots+\frac{1}{n!}\left(\Delta\boldsymbol{x}\frac{\partial}{\partial\boldsymbol{x}}+\Delta\boldsymbol{y}\frac{\partial}{\partial\boldsymbol{y}}\right)^n f(\boldsymbol{x}_0,\boldsymbol{y}_0)$$

其中

$$\left(\Delta\boldsymbol{x}\frac{\partial}{\partial\boldsymbol{x}}+\Delta\boldsymbol{y}\frac{\partial}{\partial\boldsymbol{y}}\right)^n f(\boldsymbol{x}_0,\boldsymbol{y}_0)=\sum_{i=0}^{n}\boldsymbol{C}_n^i\Delta\boldsymbol{x}^i\Delta y^{n-i}\frac{\partial^n \boldsymbol{f}}{\partial^i\boldsymbol{x}\partial^{n-i}\boldsymbol{y}}\bigg|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$$

参 考 文 献

[1] Wiener N. Extrapolation, interpolation, and smoothing of stationary time series, with engineering applications[M]. The United States: Technology Press of the Massachusetts Institute of Technology, 1950.

[2] Lauritzen S L. Time Series Analysis in 1880: A Discussion of Contributions Made by T. N. Thiele [J]. International Statistical Review, 1981, 49(3): 319-331.

[3] Lauritzen S L. Thiele-Pioneer in Statistics[M]. Oxford: Clarendon Press, 2002.

[4] Grewal M S, Andrews A P. Applications of Kalman Filtering in Aerospace 1960 to the Present [Historical Perspectives][J]. Control Systems IEEE, 2010, 30(3): 69-78.

第2章 线性系统的卡尔曼滤波

2.1 估计和最优估计

2.1.1 估计

概念:从带有随机噪声的观测数据中提取有用信息。

数学描述:假设被估计量 $\boldsymbol{x}(t)$ 为一个 n 维向量,$\boldsymbol{z}(t)$ 为 m 维观测向量,$\boldsymbol{v}(t)$ 为观测误差,$\boldsymbol{z}(t)=h(\boldsymbol{x}(t),\boldsymbol{v}(t),t)$,则估计就是利用 $\boldsymbol{z}=\{\boldsymbol{z}(\tau),t_0\leqslant\tau\leqslant t\}$ 构造一个 $\boldsymbol{z}$ 的函数 $\hat{\boldsymbol{x}}(\boldsymbol{z})$ 去估计 $\boldsymbol{x}(t)$ 的问题,则称 $\hat{\boldsymbol{x}}(\boldsymbol{z})$ 为 $\boldsymbol{x}(t)$ 的估计值,或 $\boldsymbol{x}(t)$ 的估计值为 $\hat{\boldsymbol{x}}(\boldsymbol{z})$ 。

估计准则:评估 $\hat{\boldsymbol{x}}(\boldsymbol{z})$ 的好坏的标准称为估计准则。

常用的估计准则:

(1) 最小二乘估计:使残差的平方和最小,即

$$\min J=(\boldsymbol{z}-\boldsymbol{H}\hat{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{z}-\boldsymbol{H}\hat{\boldsymbol{x}}) \tag{2-1}$$

(2) 最小方差估计:使估计误差方差阵达到最小,即

$$\begin{aligned}\min J&=E_{x,z}(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\hat{\boldsymbol{x}})\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\hat{\boldsymbol{x}})p(\boldsymbol{x},\boldsymbol{z})\,\mathrm{d}x\mathrm{d}z\end{aligned} \tag{2-2}$$

(3) 极大似然估计:使条件概率密度 $p(\boldsymbol{z}\mid\boldsymbol{x})$ 达到极大,即

$$\max J=p(\boldsymbol{z}\mid\hat{\boldsymbol{x}}) \tag{2-3}$$

(4) 极大验后估计:使验后概率密度 $p(\boldsymbol{x}\mid\boldsymbol{z})$ 达到极大,即

$$\max J=p(\hat{\boldsymbol{x}}\mid\boldsymbol{z}) \tag{2-4}$$

2.1.2 最小二乘估计

当不知道 $\boldsymbol{x}$ 和 z 的概率分布密度，也不知道它们的一、二阶矩阵时，就只能采取高斯提出的最小二乘法进行估计。最小二乘估计是高斯在 1795 年为测定行星轨道而提出的参数估计算法。最小二乘估计是以残差的平方和为最小作为估计准则的。这种估计的特点是算法简单，不必知道与被估计量及量测量有关的任何统计信息。

2.1.2.1 经典最小二乘估计

设被估计量 $\boldsymbol{x}$ 是 n 维随机向量，为了得到其估计，如果对它进行 k 次线性观测，得到

$$z_i = \boldsymbol{H}_i\boldsymbol{x} + \boldsymbol{v}_i \quad (i = 1,2,\cdots,k) \tag{2-5}$$

式中：z_i 为 m 维观测向量；$\boldsymbol{H}_i$ 为 $m \times n$ 观测矩阵；$\boldsymbol{v}_i$ 为均值为零的 m 维观测误差向量。最小二乘估计是使 $J = \sum_{i=1}^{k} (z_i - \boldsymbol{H}_i\hat{\boldsymbol{x}})^{\mathrm{T}}(z_i - \boldsymbol{H}_i\hat{\boldsymbol{x}})$ 最小。

令 $z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix}$，$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \\ \vdots \\ \boldsymbol{H}_k \end{bmatrix}$，$\boldsymbol{v} = \begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_k \end{bmatrix}$，则 $J = (z - \boldsymbol{H}\hat{\boldsymbol{x}})^{\mathrm{T}}(z - \boldsymbol{H}\hat{\boldsymbol{x}})$。

为使 $J = (z - \boldsymbol{H}\hat{\boldsymbol{x}})^{\mathrm{T}}(z - \boldsymbol{H}\hat{\boldsymbol{x}})$ 最小，对 J 求偏导，并令其等于零，即

$$\begin{aligned}
\frac{\partial J}{\partial \hat{\boldsymbol{x}}} &= \frac{\partial\,(z - \boldsymbol{H}\hat{\boldsymbol{x}})^{\mathrm{T}}(z - \boldsymbol{H}\hat{\boldsymbol{x}})}{\partial \hat{\boldsymbol{x}}} \\
&= \frac{\partial(z^{\mathrm{T}}z - z^{\mathrm{T}}\boldsymbol{H}\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}z + \hat{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\hat{\boldsymbol{x}})}{\partial \hat{\boldsymbol{x}}} \\
&= -z^{\mathrm{T}}\boldsymbol{H} - z^{\mathrm{T}}\boldsymbol{H} + 2\hat{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} \\
&= -2\boldsymbol{H}^{\mathrm{T}}(z - \boldsymbol{H}\hat{\boldsymbol{x}}) \\
&= 0
\end{aligned} \tag{2-6}$$

由式(2-6)可得 $\boldsymbol{H}^{\mathrm{T}}z = \boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\hat{\boldsymbol{x}}$。所以最小二乘估计的估计值为

$$\hat{\boldsymbol{x}} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}z \tag{2-7}$$

最小二乘估计的几点说明：

(1) 最小二乘估计需要把所有的观测量一起处理，属于批处理过程，不是实时估计，但优点是不需知道任何先验信息。

(2) 最小二乘估计是无偏估计。