Bayesian Filtering and Smoothing

# 贝叶斯滤波与平滑

[芬] 希莫•萨日伽（Simo Särkkä） 著
程建华 陈岱岱 管冬雪 等译
杨晓东 袁书明 审校

国防工业出版社
National Defense Industry Press
CAMBRIDGE

国家自然科学基金资助(61374007、61273081、61104036)

# Bayesian Filtering and Smoothing

# 贝叶斯滤波与平滑

[芬] 希莫·萨日伽(Simo Särkkä) 著
程建华 陈岱岱 管冬雪 等译
杨晓东 袁书明 审校



国防工業出版社
·北京·

**著作权合同登记　图字:军-2015-082号**

**图书在版编目(CIP)数据**

贝叶斯滤波与平滑/(芬)萨日伽著;程建华等译.
—北京:国防工业出版社,2015.8
书名原文:Bayesian Filtering and Smoothing
ISBN 978-7-118-10247-5

Ⅰ.①贝...　Ⅱ.①萨...②程...　Ⅲ.①贝叶斯理论-滤波理论
②贝叶斯理论-平滑　Ⅳ.①TP13 ②TN911.23

中国版本图书馆 CIP 数据核字(2015)第171928号

This is a translation of the first title published by Cambridge University Press:
Bayesian Filtering and Smoothing by Simo Särkkä
ISBN 978-1-107-61928-9
Cambridge University Press 2013
This translation for the People's Republic of China (excluding Hong Kong, Macau and Taiwan) is published by arrangement with the Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.


※

国防工業出版社 出版发行

(北京市海淀区紫竹院南路23号　邮政编码100048)
北京嘉恒彩色印刷有限责任公司
新华书店经售

*

**开本** 710×1000　1/16　**印张** 12¼　**字数** 225千字
2015年8月第1版第1次印刷　**印数** 1—2000册　**定价** 45.00元

**(本书如有印装错误,我社负责调换)**

国防书店:(010)88540777　　发行邮购:(010)88540776
发行传真:(010)88540755　　发行业务:(010)88540717

# 译者序

滤波与平滑是估计理论中最为核心的两类算法，可用于估计未知的状态或参数。贝叶斯滤波与平滑，指的是贝叶斯意义下的滤波与平滑，其内容涵盖了经典线性滤波与平滑算法，非线性滤波与平滑算法，以及非高斯滤波与平滑算法。作为统计学的分支，贝叶斯学派依托经典的贝叶斯公式和推理，发展了有别于传统频率学派的理论与方法。随着信息科学的迅猛发展，作为强有力的数据处理算法，贝叶斯滤波与平滑算法在诸多领域都得到了极为广泛的应用，包括自动控制、导航、金融、通信、航空、航天、医疗等。

芬兰阿尔托大学 Simo Särkkä 博士长期从事贝叶斯推理方法及其应用相关的研究工作，在目标追踪、机器学习、脑成像等方面取得了卓著的研究成果。2013 年，作者通过整理在芬兰赫尔辛基科技大学、阿尔托大学以及坦佩雷科技大学的上课讲义，以及在英国剑桥大学做访问学者的研究成果，出版了学术著作 *Bayesian Filtering and Smoothing*，详细阐述了线性/非线性滤波算法，同时还特别针对新型粒子滤波、线性/非线性平滑算法和参数估计方法进行了系统的阐述。该著作风格简洁，理论严谨，体系完善，算法先进，贴近工程实用。Simo Särkkä 博士不仅是贝叶斯估计领域优秀的学者，还拥有丰富的工程经验。相信著作的引进，能为国内相关院校和科研院所在贝叶斯滤波与平滑方面的研究提供有益的参考。

为了更好地学习和分享先进的学术成果，我们组织翻译了《贝叶斯滤波与平滑》一书。本书的翻译和审定工作由哈尔滨工程大学程建华副教授负责。其中，前言、第 5 章、第 6 章、第 7 章、第 8 章和第 9 章由程建华完成；第 10 章、第 11 章、第 12 章、第 13 章和附录由陈岱岱博士完成；第 1 章和第 3 章由管冬雪博士完成；第 2 章由齐兵博士完成；第 4 章由王通达博士完成。

在本书的翻译和出版过程中，海军潜艇学院杨晓东教授和海军装备研究院袁书明高工对书稿进行了认真、细致的审阅，并提出了许多宝贵的意见和建议。在书稿翻译期间，陈岱岱博士获得国家留学基金委资助，对原著作者所在的阿尔托大学贝叶斯课题组进行了为期一年的访问交流，Simo Särkkä、陈芳娟、陈新桃、Arno Solin 和 Songzier Tiny 对著作的翻译给予了大力支持和协助。在此表示诚挚的谢意！

感谢实验室董金鲁、王冰玉、屈传波、李美玲、徐英蛟、孙湘钰、宋春雨、于天

琦、刘萍和张昊等硕士研究生对本书初期翻译和后期校对工作的支持与帮助。

感谢国家自然科学基金(61374007、61273081、61104036)、中央高校基本科研业务费(HEUCFX41309)和参译人员所在单位对本书出版的支持。

感谢国防工业出版社曲岩编辑对本书的版权引进和译著出版给予的支持和帮助。

由于译者水平所限,翻译过程中难免存在不妥之处或错误,敬请读者和专家们批评指正。

译 者

2015 年 2 月

# 前　言

本书主要论述了非线性卡尔曼滤波及平滑，粒子滤波及平滑，以及相关的参数估计方法。尽管力求简明、扼要，但书中所有估计方法均会给出严谨的数学解释。所以具备一定数学基础的读者，均能较好地理解所述的估计方法。本书篇幅力求精短，以方便读者的快速阅读。

本书主要面向与应用数学和计算科学专业相关的高年级本科生及研究生。对于需要应用相关估计方法的研究人员和工程师也同样适用。本书需要读者具备一些基本的相关知识，主要包括线性代数、矢量计算、贝叶斯推理以及 MATLAB 编程。

如书名所述，书中所涉及的模型与算法均设定在贝叶斯意义下，即所有估计的结果均可视为对特定概率分布或相应参数的逼近。概率分布通常用于表示系统模型中的不确定性，以及对随机物理特性的建模表征。书中所涉及的非线性滤波、平滑，以及参数估计方法均将以贝叶斯推理形式表述。基于相同的贝叶斯符号与公式，还将推导经典及新型算法。书中涉及的贝叶斯方法可追溯至 20 世纪五六十年代 Stratonovich 的先驱性工作，这在时间上甚至比 1960 年卡尔曼提出经典卡尔曼滤波还要早。因此，非线性滤波理论从一开始就符合贝叶斯意义。

第一章大致介绍了贝叶斯滤波与平滑的概念与应用。第二章简要地综述了贝叶斯推理的基本内容和进行贝叶斯解算所涉及的数学基本方法。第三章通过结合用递归方法解决线性回归问题的例子，首先逐步推导了贝叶斯估计方法。通过扩展和推广该问题，引出了贝叶斯滤波与平滑理论。该章还介绍了书中的第一种卡尔曼滤波器。

第四章通过推导贝叶斯滤波方程及著名的卡尔曼滤波器，开始正式论述贝叶斯滤波理论。第五章着重介绍了卡尔曼滤波的几种非线性扩展，包括扩展卡尔曼滤波、统计线性化滤波和无迹卡尔曼滤波。第六章将上述几种滤波统一到了广义高斯滤波框架下，并由该滤波框架推导出了高斯 – 埃尔米特卡尔曼滤波以及容积卡尔曼滤波。在第七章中，介绍了基于序列蒙特卡洛方法的粒子滤波，首先介绍了基本 SIR 滤波器，最后介绍了 Rao – Blackwellized 粒子滤波器。

第八章首先推导了基本常规（固定区间）贝叶斯平滑方程，然后推导了其中

的特例 RTS 平滑器。第九章介绍了扩展 RTS 平滑器,统计线性化 RTS 平滑器以及无迹 RTS 平滑器。第十章中给出了广义高斯平滑框架,并推导了其中的两种特例:高斯 - 埃尔米特 RTS 平滑器和容积 RTS 平滑器。此外在这一章中,还讨论了高斯固定点及固定滞后平滑。在第十一章中,首先利用改进的基本 SIR 粒子滤波器实现类似平滑的功能。接着,介绍了数值性更优的后向仿真粒子平滑器,以及变加权(或边缘)粒子平滑器。最后,讨论了 Rao - Blackwellized 粒子平滑器。

第十二章介绍了状态空间模型的参数估计方法,内容主要集中在以下三部分:估计的最优化,基于期望极大化准则的极大似然估计与极大后验估计,以及马尔可夫链蒙特卡洛方法。我们先提出大概的参数估计方法,再详细阐述在进行参数估计时如何利用所述的卡尔曼滤波器及其 RTS 平滑器、非线性高斯滤波器及其 RTS 平滑器、粒子滤波器与平滑器,去计算或逼近所需的参数估计值。例如,基于经典的状态空间模型期望极大化算法,可以推导出粒子期望极大化算法或者粒子马尔可夫链蒙特卡洛算法。此外,我们还讨论了 Rao - Blackwell 化在某些条件下对参数估计的有益辅助。

第十三章作为本书的末章,我们就不同应用对象如何选取不同的估计方法,给出了一些通用的建议。我们也对多个关键的技术进行了讨论,并给出了相关的参考文献。这些参考文献的内容很重要,但是并没有写进书中。

在每一章节的最后,都附有一些习题。便于读者进行实践,提升读者在估计算法的实现以及就具体问题选用合适的估计方法方面的经验。习题所对应的 MATLAB 源代码以及相关的资料,均可在本书的网址进行下载。本书的网页链接为 www. cambridge. org/sarkka。

本书的内容源自于作者本人于 2009 年至 2012 年期间在芬兰赫尔辛基科技大学、阿尔托大学以及坦佩雷科技大学上课时所用的讲义。书中大部分的内容是我在阿尔托大学(原赫尔辛基科技大学)生物医学工程及计算科学系工作时完成的,还有一部分内容是我在英国剑桥大学工程系做访问学者时撰写的。感谢芬兰科学院复杂计算系统研究中心、生物医学工程及计算科学系与阿尔托大学理学院提供的研究基金,本书才得以顺利出版。

感谢生物医学工程及计算科学系的 Jouko Lampinen 教授和 Aki Vehtari 教授,给我提供研究与合作的机会,这本书最终才得以产生。感谢 Arno Solin, Robert Piché, Juha Sarmavuori, Thomas Schön, Pete Bunch 和 Isambi S. Mbalawata 仔细校对了书稿,并提出了很多有益的建议。非常感谢我的研究伙伴 Jouni Hartikainen, Ville Väänänen, Heikki Haario 和 Simon Godsill,提升了我在这一领域的知识,并发展了部分书中提及的方法。感谢剑桥出版社的 Diana Gillooly 以及高

级编辑 Susan Holmes，是他们建议我对上课讲义进行整理和出版。最后，我还要感谢我妻子 Susanne 在书稿写作期间的支持与宽容。

Simo Särkkä
万塔，芬兰

# 目 录

# 第一章　贝叶斯滤波与平滑

最优滤波,通常指的是用于估计时变系统状态的一类方法。在时变系统中,状态可通过含有噪声的量测量进行间接观测。此处的“最优”指的是统计意义下的最优。贝叶斯滤波就是指贝叶斯意义下的最优滤波。本书所涉及的滤波均为贝叶斯滤波。

在最优滤波、贝叶斯滤波和贝叶斯最优滤波中,系统状态指的是位置、速度、方向和角速度等动态变量的集合,能够充分描述相应的系统。量测量中的噪声标示着量测量的不确定性;即使真实的系统状态已知,量测量也无法与系统状态共同满足一个确定的函数关系,只能得到可能的量测值的一个分布。通常,状态变量的时间演变特性用动态系统模型进行表示。动态系统由一个确定的过程噪声驱动,该噪声被用于表示动态系统中不确定性。尽管在大多数情况下,系统并非真正的随机,但随机特性还是常被用于表示系统模型的不确定性。

贝叶斯平滑(或最优平滑)通常被认为是贝叶斯滤波范畴中的一类方法。贝叶斯滤波的基本形式是利用过去的量测量计算当前系统状态的估计值,而贝叶斯平滑则可用于重构当前时刻之前的系统状态。虽然在一般意义上,“平滑”通常表示产生一种使数据更加平滑的方法(而不是粗糙的)。而在贝叶斯滤波的内容中,“平滑”一词有着更为明确的含义。

## 1.1　贝叶斯滤波和平滑的应用

在许多工程应用领域中,通常可以建立起上述系统时变模型,例如:导航、航空工程、航天工程、远程监测、通信、物理、音频信号处理、控制工程、金融等。

- 全球定位系统(Global Positioning System, GPS)[83]是一种被广泛应用的卫星导航系统,通过 GPS 接收单元测定多个 GPS 卫星信号的波达时间,并利用这些测量数据计算出接收机的位置,如图 1.1 所示。通常使用扩展卡尔曼滤波器(Extended Kalman Filter, EKF)或其他最优滤波算法①计算 GPS 接收机当前的位置和速度,因此需要考虑量测量和设定的动态模型(符合物理定律)。同时,卫星星历作为卫星向 GPS 接收机发射的参考信息,也通常采用最优滤波进行解算。

① 严格地说,扩展卡尔曼滤波器仅为最优滤波的近似算法,因为它使用基于泰勒级数的高斯滤波算法来逼近非高斯最优滤波算法。

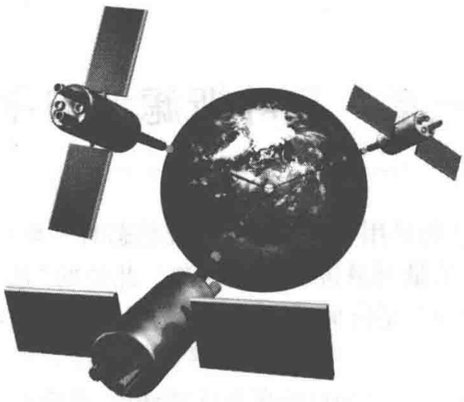

图 1.1　在 GPS 系统中,量测信息包含了卫星信号的时间延迟,通常使用最优滤波器(如扩展卡尔曼滤波器)来计算位置和准确的时间。

• 目标跟踪[13,24,27],是指利用多种/个传感器来确定远程目标位置和速度的方法,如图 1.2 所示。典型的传感器有有源雷达、无源雷达、射频传感器、声学阵列、红外传感器等。当对目标进行连续追踪时,利用最优滤波或平滑算法,结合目标的动态方程以及各传感器的量测量进行解算。在单目标跟踪中,目标可以针对一个机器人,一颗卫星,一辆汽车或一架飞机等。

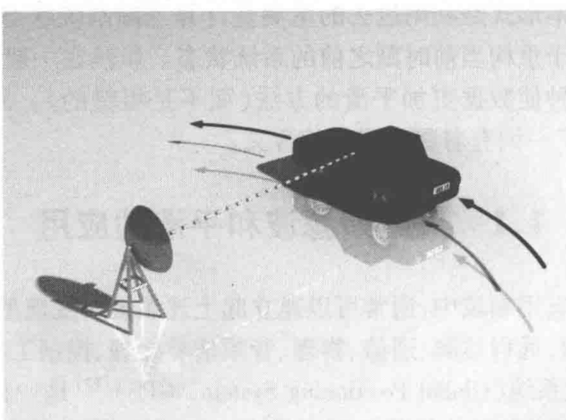

图 1.2　在目标跟踪中,利用传感器(如雷达)产生目标的量测信息(如方位和距离),用于确定目标轨迹。

• 多目标跟踪[12,20,144,160]系统,可用于同时实现对某一地区内多个目标的远程监控,如图 1.3 所示。这将涉及数据关联和目标数量的估计问题。其中,数据关联指的是要解决量测量来自哪个目标的问题。多目标跟踪系统通常用于军事领域中的远程监控,但在民用领域也有诸多应用,如汽车隧道监测、自动报警

系统以及室内行人追踪等。

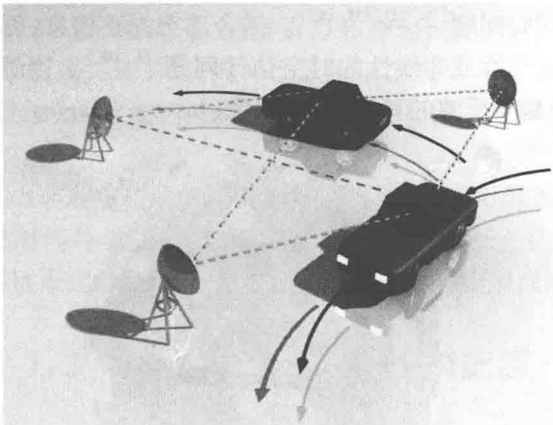

图 1.3　在多目标跟踪中,数据关联问题必须得以解决,因为在没有任何辅助信息的条件下,无法判断量测信息来自哪个目标。

• 惯性导航[56, 166],利用加速度计和陀螺仪等惯性传感器所敏感到的运动信息,计算汽车、飞机或导弹等载体的位置和速度。考虑到传感器的测量误差(存在量测噪声),通常使用最优滤波或平滑计算估计值。另外,最优滤波和平滑也常被应用于时变环境中传感器的校准。

• 组合惯性导航[13, 56],结合了两类传感器的优势:一种是长时间工作误差不积累但量测不够精确的传感器,如高度计和地标跟踪器等;另一种是误差累积但局部/短时量测精准的惯性传感器。通常利用扩展卡尔曼滤波器等最优滤波器,对来自于不同系统的信息进行融合。例如,1969 年登陆月球的"阿波罗"11 号登月舱(鹰),其制导系统就应用了组合导航技术。

• GPS / INS 导航[13, 56],是组合惯性导航的一种:惯性导航系统(Inertial Navigation System, INS)与 GPS 接收机单元进行组合导航。在 GPS/ INS 组合导航系统中,GPS 的短期扰动可由惯性传感器进行补偿,同时惯性传感器的误差可由 GPS 接收机进行补偿。这种方法的另一个优点是,当 GPS 接收机因为某些原因无法计算其所处位置(如未能捕获卫星信号)时,导航系统可以暂时切换至纯惯性导航模式下进行工作。例如,在室内、隧道及其他情况下,GPS 接收机和卫星之间受遮挡时,通常就采用这种方法。

• 脑部成像方法,通过利用最小范数估计(Minimum Norm Estimates, MNE)及其推广形式[62, 78, 93, 165],结合由传感器输出含噪声的数据重构大脑释放上述信息的源区域。如脑电图(Electroencephalography, EEG)、脑磁图(Magnetoencephalo - graphy, MEG)、并联功能磁共振成像(Functional Magnetic Resonance Imaging, fMRI)和扩散光学层析成像(Diffuse Optical Tomography, DOT),如图 1.4 所示。就贝叶斯角度而言,最小范数解也可以理解为利用高斯观测量去估

计具有特定结构的区域。通过上述理解方式,估计问题将转化为统计反演问题或广义高斯过程回归问题[78, 122, 135, 165]。结合动态先验信息,可以用卡尔曼滤波和平滑来解决这些线性或非线性的时空估计问题[65,146]。当然,对于并行 fMRI 问题,也可以采用基于反演的算法,如逆成像(Inverse Imaging, InI)方法[93]。

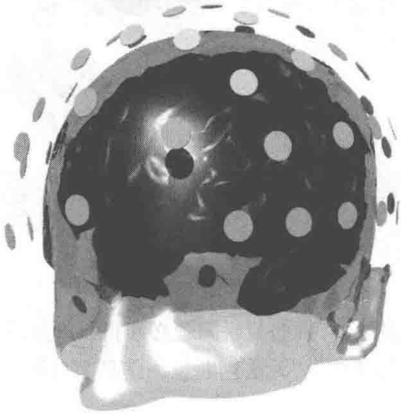

图 1.4　脑部成像方法利用传感器输出数据来估计脑部状态,如 EEG 和 MEG。在动态情况下,相关的反演问题可以用最优滤波或平滑方法解决。

- 感染性疾病的传播[84],通常可以建模为关于易感人群、感染人群、恢复人群及死亡人群的微分方程。当动态方程具有不确定性且量测信息不精确时,疾病传播的估计问题可以表述为最优滤波问题[142]。
- 生物过程[106],如人口增长、捕食模型及生物学其他动态过程,也可以建模为(随机)微分方程。利用不准确的测量信息进行过程状态估计,也可以表述为最优滤波和平滑问题。
- 通信,也是最优滤波的传统应用领域。例如,最优接收器、信号检测器以及锁相环等[119, 171, 172]。对于基本的隐马尔可夫模型(Hidden Markov Model, HMM),著名的 Viterbi 算法[174]也可视为一种用于计算最大后验(Maximum A Posteriori, MAP)贝叶斯平滑解的方法。
- 音频信号处理,如音频恢复[48]和音频信号增强[42]等应用中,通常利用时变自回归(Time - Varying Autoregressive, TVAR)模型作为基本的音频信号模型。利用最优滤波和平滑,这些模型能够得到有效的估计。
- 随机最优控制[98, 159],是指对时变随机系统的控制。随机控制器在飞机、汽车和火箭等控制领域中具有典型的应用。除了统计最优性,最优还意味着控制信号在达到一定性能的同时所需的代价最小,如达到预定状态所需的期望时间最小、消耗燃料最少或离期望位置轨迹的平均距离最短等。传感器对系统状态的观测,通常需要利用最优滤波处理传感器获得的量测数据,对系统状态进行重构。

● 学习系统或自适应系统在数学上通常可以表示为最优滤波和平滑问题[64]，并且与贝叶斯无参数建模、机器学习和神经网络建模[19]紧密关联。类似于多目标跟踪中数据关联方法，同样也适用于在线自适应分类[4]。高斯过程回归[122]和最优滤波之间的联系，也已在文献[61]、[138]与[143]中进行了论述。

● 时变的且由非理想传感器量测的物理系统，有时可建模为随机状态空间模型，并且系统的时间特性可以用最优滤波进行估计[78]。这类问题通常被称为逆问题[165]，而且最优滤波和平滑通常被视为解决时变逆问题的贝叶斯方法。

## 1.2 贝叶斯滤波和平滑的起源

时间相关问题的贝叶斯分析方法，其本质属于最优线性滤波范畴。由于线性系统的数学表述简易，因此最早针对线性系统构建最优递归估计器的思想。无论从数学角度还是从建模角度来看，最自然的最优性准则是最小二乘。对于线性系统，最优贝叶斯方法（均方误差最小（Minimum Mean Squared Error，MMSE））与最小二乘法等价，即最优最小二乘解就是后验均值。

最优滤波最早可追溯到维纳滤波[177]，维纳滤波是在频域中解决平稳高斯信号的最小二乘最优滤波问题的方法。至今为止，维纳滤波在通信[119]、数字信号处理[63]及图像处理[52]等领域内仍十分重要，但维纳滤波的缺点在于只适用于平稳信号。

最优线性滤波在工程应用中的成功，主要归功于卡尔曼的开创性论文[80]，该论文阐述了最优离散（采样）线性滤波问题的递归解。相比于维纳滤波，卡尔曼滤波具有更简单的数学结构，便于被更好地理解与应用，因此大获成功。然而，尽管卡尔曼滤波更为简易通用，但卡尔曼滤波（或更精确地说是 Kalman – Bucy 滤波器[81]）仍包含了维纳滤波，并视维纳滤波为它的一个特例。

在其发展的早期阶段，卡尔曼滤波很快被发现属于贝叶斯滤波范畴[66, 71, 72, 90]。在卡尔曼滤波出现后不久，相应的贝叶斯平滑[92, 123, 124]也得到了发展。一个有趣的历史事实是，当 Kalman 和 Bucy 在美国发展这一线性理论时，Stratonovich 也正在俄罗斯对概率（贝叶斯）方法进行着开创性的研究[72, 162]。

正如文献[176]所述，贝叶斯学派的学者也应用了类似于递归估计器的卡尔曼滤波器，而且尚不清楚卡尔曼滤波理论和动态线性模型（Dynamic Linear Models，DLM）理论哪个出现在前。尽管上述两种理论最初的出发点略有不同，但二者是等价的。由于卡尔曼滤波与随机最优控制理论及其发展具有密切的联系，本书将从卡尔曼滤波展开对贝叶斯滤波问题的阐述。

虽然卡尔曼滤波最初的推导基于最小二乘法，但是从纯概率贝叶斯分析中