估计理论及其在通讯与控制中的应用

〔美〕 A. P. 塞奇　J. L. 梅尔萨 著

科 学 出 版 社

估计理论及其在通讯与控制中的应用

〔美〕 A. P. 塞 奇
J. L. 梅尔萨 著

田承骏 唐策善 等译
潘一民 常学将 校

科 学 出 版 社

1 9 7 8

内 容 简 介

本书为介绍估计理论及其在通讯与控制中的应用的专著. 全书共九章, 第一章导论; 第二至第四章为概率论与随机过程的基础知识; 第五章为判定理论; 第六章为估计的基本理论; 第七至第九章为滤波理论, 分别介绍最优线性滤波、最优线性滤波的推广及非线性估计.

本书为应用数学方面的著作, 偏重于应用, 书中的各种滤波算法均列成表格, 便于使用.

读者对象为从事通讯与控制的工程技术人员和数学工作者.

A. P. Sage　J. L. Melsa

ESTIMATION THEORY

with Applications to Communications

and Control

McGraw-Hill, 1971

估计理论及其在通讯与控制中的应用

〔美〕A. P. 塞　奇　J. L. 梅尔萨　著

田承骏　唐策善　等译

*

科学出版社 出版

北京朝阳門内大街137号

北京印刷二厂印刷

新华书店北京发行所发行　各地新华书店经售

*

1978年5月第一版　开本: 787×1092 1/32

1978年5月第一次印刷　印张: 19 3/4

印数: 0,001—18,100　字数: 452,000

统一书号: 13031・741

本社书号: 1067・13—1

定　价: 2.00元

中 译 本 序

Sage 和 Melsa 为从事通信及控制工作者所写的这本估计理论，与当前同类的书相比较，是一本宜于自学的读物，因为读它时需要的数学预备知识较少．它着重讲状态空间法，这对了解近代滤波及控制理论是必不可少的方法．根据毛主席所教导的"**古为今用，洋为中用**"的方针，把它译成汉文，可以为国内从事这方面工作的同志提供参考．

书中第一章介绍了全书的概貌．第二、三章对本书所要用的概率方面的知识，给以扼要的复习．但读者如果对概率及随机过程以前毫无接触，凭这本书所介绍的内容，还不能得到足够的认识，从而也不可能顺利地接受其余各章的内容．

第四章主要内容是讲扩散型随机微分方程，这一方程是现代随机控制论中用来描写被控系统的数学模型．在这一章里，为了讲扩散型随机微分方程，事先直观地介绍了 Wiener 过程、随机积分及伊藤微分法则等概念，这为初次遇到这些概念的读者们，提供一个较易接受的讲法．

第五章介绍统计判决理论．作为应用，在最后一节里讲 Gauss 噪声干扰下检测马氏信号的问题，通过它可以看出判决理论与估计理论之间的密切联系．

第六、七、八章是本书的重点，它系统地介绍了现代应用很广的滤波理论．第六章系统地介绍了点估计理论．首先讲 Bayes 估计理论，然后逐渐放宽关于验前统计知识，考虑极大验后估计和极大似然估计，并讨论它们的性质．本章第五节讨论当验前均值及方差不对时，它对极大似然估计和极大验

后估计所产生的误差进行了分析，这种分析是针对通信及控制的应用而提出的．本章最后两节讲线性最小方差估计和工程技术中常用的最小二乘估计，前者只对验前统计知识假定其前两阶矩存在，而后者则对所有的验前统计知识不加任何假定，换言之即不考虑其统计性质．在这一章里给了一些例子，其中有些是联系着通信及控制的实际而拟定的．

第七章讲 Kalman-Bucy 线性滤波，它用两种不同方法导出离散时间的滤波公式．对连续时间的滤波公式，除了用离散时间的公式取其极限方法推导外，还讲了变分法方法及 Wiener-Hopf 方法．通过这些不同方法的推导，可以加深了解滤波公式的物理及统计的要点．本章最后一节讲滤波的稳定性，但它只陈述了一些结果，基本未给证明，由于这一问题在滤波应用中很重要，希望读者参阅书中所列的参考文献，或参考《离散时间系统滤波的数学方法》(中国科学院数学研究所概率组，国防工业出版社)．

第八章把前一章的结果推广到有色量测噪声的情形．对于连续时间的平滑问题，采用了新息方法推导其公式，新息的概念在滤波问题中是一个很重要的概念．本章最后一节讨论了实际应用中，当模型取得不准时的误差分析问题，关于这一问题的一般的讲法，当可参看《离散时间系统滤波的数学方法》．

第九章讨论非线性滤波，对于一般的非线性滤波方程的推导需要较深的数学知识，本书所采用的直观的办法使初学者容易接受．本章其后各节是对一般的非线性滤波公式，进行一阶及二阶的近似，以便求得可行的计算．对求极大验后滤波的近似算法，本章还采用了不变嵌入原理方法．

总之，本书对初学估计理论及其在通信及控制方面的应用来说，是较比合适的读物，通过它可以得到关于近代滤波理

序 言

在通讯与控制领域内，目前十分重视估计理论的研究．本书对估计理论及其在通讯与控制中的应用作一全面论述，可适合于这些学科的研究生阅读．本书是由工程技术人员为工程技术界而写的，因此，为了描述基本概念并阐明各种方法的相似性和局限性，我们基本上不考虑纯数学的严格性问题．

信息与控制科学中心的几个不同的课程曾使用过本书．最初用于研究生的估计理论课，他们曾学过随机过程（采用 Papoulis 著《概率、随机变量与随机过程》或其它类似的教材）．许多学生还学过最优化理论（以 Sage 著《最优系统控制》或 Schultz 和 Melsa 合著的《状态函数与线性控制系统》为教材）．他们一学期的课程内容包括第四章的后半部分和第五至第九章． 经验还表明，对于不熟悉随机过程但已初步掌握概率论的学生来说，本书是很有用的．在这种情况下，学习的内容应包括第二、三章和第四章的前半部分，以及第六、七章的部分内容．然后再学习第四章的后半部分和第五、八、九章；这样，也许要多安排一个学期．当然，为适合不同学生的基础与兴趣，这两种方案可能有相当大的变动．

（下略）

A. P. 塞　奇
J. L. 梅尔萨

目　录

第一章　导　论

工程技术人员和许多人所面临的重要任务之一，就是如何很好地使用观测中的信息去作出判定．本书将介绍用以作出判定的各种数学方法，以便在用全部可以利用的信息来权衡（或影响）时去确定最优判定．这类数学方法统称为估计．它和信息一词有着内在的联系，我们将认为信息是作出判定的有用数据．

上述问题可以说是一个老问题，至少可追溯到 Legendre (1806) 与 Gauss (1809)．高斯 (Gauss) 被认为是首先使用估计方法的(用于轨道确定)，根据他的意见，被估计参数的最可能值是这样的，它使各实际观测和计算值的差的平方乘上表示该观测值相对置信程度的加权系数的和为最小．这就是最小二乘估计的概念，我们将在第六和第九章中详加讨论．

我们所讨论的估计理论联系到许多不同的领域．通讯理论大部分涉及到根据接收信号知识去确定发送消息的特征的问题，这些接收信号是发送消息的调制形式受噪声污染的观测结果．控制论研究自动系统的设计问题，这个系统根据对周围环境的观测结果确定出施加于一个装置的控制，或者确定所施加的控制以抵消由一个精明的对手所加控制的影响*．运筹学或工业工程技术与资源分配、清点存货控制、编制计划等等判定问题有关．

为把估计理论的方法应用于一个专门的工程技术问题，

* 例如，火力控制系统应能作出必要的控制以对付目标的机动性．——译者注

必须把这一问题用适当的数学语言表达出来；这就是概率论与随机过程的语言．在学习了估计理论所必需的概率论与随机过程的有关内容之后，我们拟深入研究估计理论的方法及其在通讯与控制的有关问题中的应用．对于这些问题推导出各种有效的算法，是本书贯彻始终的目标．具体地说，各章的主要内容如下：

第二章　概率与随机变量　概率论是研究经验事件的平均性质的．经验知识的任何一种数值度量都是由一些命题集组成的，这些命题就它们的正确性而论各有不同的可信程度．用数值表示这种程度，其有关课题就是概率论的研究对象．虽然我们希望读者已经掌握了概率论，但仍在第二章对概率论作一个简短的回顾，以使读者熟悉在本书其余部分所使用的各种符号．我们主要是研究期望，因为正如我们将要看到的那样，随机变量的条件期望常常产生出我们所要的最优估计量．我们也分析了验前密度和验后(或条件)概率密度的差别．

第三章　随机过程　在第三章中考虑作为时间函数的随机过程概念，并对它进行了研究．我们引入重要的协方差和相关函数的概念，它们是随机过程的时间相关性的度量．最后，我们主要致力于研究线性连续系统与线性离散系统对于随机过程输入的响应．

第四章　高斯-马尔柯夫过程与随机微分方程　通常，仅当某些统计量具有高斯分布时，才能导出计算上有效的估计算法．由于高斯过程对于我们后面的研究起着重要作用，所以本章对它做了详细讨论．我们首先讨论中心极限定理，它确立了这样的事实，(幸好)许多随机过程至少是近似高斯的．

对于一大类问题来说，马尔柯夫(Марков)假定也是成立的，它在下文的推导中将很有用．马尔柯夫假定说明，现在的

知识把过去同将来分开．这个假定赋予在第二章中引入的条件期望概念以特殊的意义．首先，我们略为详细地分析线性系统中的高斯-马尔柯夫过程，然后讨论非线性系统中的随机过程．

在研究非线性系统时，我们会发现，通常的微积分运算法则并不是总能用的．我们将介绍 Ito 随机微积分，它可以解决涉及到非线性系统中随机过程的许多基本问题．Fokker-Planck 偏微分方程的导出和它的近似解将描述出连续非线性系统中的均值与方差传播．对于高斯随机变量应用条件期望定理，使得我们可以讨论离散非线性系统中的均值与方差传播．

第五章　判定理论　通常认为判定理论与估计理论是有区别的，并且这两类问题构成了统计推断的两个主要领域．这里从另一个方面来说，重点在于强调估计和检测这两个概念的相似性．从第一章至第四章，大部分是讲随机信号以及线性系统和非线性系统对信号的处理．虽然我们考虑了研究嵌入在噪声中的随机过程的可能性，但从未考虑信号是否出现的问题；而这就是第五章要讨论的内容．

也许，判定理论的最重要的应用是检测由于干扰或噪声而弄模糊的信号．它要解决的问题就是，以最优方式判断是否收到了想要的信号．我们依据不同判定的相应代价，可以定出一个判定门限，用观测的组合同它比较来进行判定．嵌入噪声中的已知信号与随机信号、以及二元与 M 元判定过程等各种问题，将用固定样本容量的假设检验和序贯假设检验两种方法进行研究．最后，我们给出判定理论问题的状态变量表述，它将指出判定理论与估计理论之间的本质联系．

第六章　估计的基本理论　本章集中介绍统计推断的第二个主要领域——状态和参数估计的一些基本内容．着重介

绍所谓点估计量．而本书第七、八、九各章将用于介绍序贯估计理论．

首先，我们认为贝叶斯(Bayes)估计理论是第五章贝叶斯判定理论的逻辑推广．如果假定被估计参数没有验前统计知识，那么就得到经典的极大似然估计量．我们将介绍估计量的许多性质，并讨论伪贝叶斯估计量与误差分析．

其次，假定已知信号与噪声的第一、二阶统计矩，并且加上估计量为线性的要求．那么，确定了最优线性估计量就推导出正交投影引理与维纳-何甫(Wiener-Hopf)方程，对于估计与判定理论的研究，它们是两个最重要的关系式．

最后，完全去掉对信号与噪声的统计知识的要求，就导致我们讨论最小二乘估计．如果适当选取加权矩阵和估计模型，可以证明最小二乘估计量等价于极大似然估计量或贝叶斯估计量．

第七章　最优线性滤波　第六章的大多数估计量不是序贯的估计量．它们是非序贯的、或非递推的，因为不是边收到观测边处理，也没有边收到观测边修改估计量．第六章的许多方法是在收到全部观测后同时进行处理，以确定出最优估计量．如果边收到观测边处理，计算上就可以有很大的节约．这种序贯处理常常是估计量的很自然的需要．

在第七章中，我们介绍线性最小误差方差序贯非平稳估计问题的解，这是由卡尔曼(Kalman)、布西(Bucy)等人推广维纳(Wiener)原始工作的结果．首先讨论这个问题的离散时间形式，以第六章的投影引理作为求解的工具．其次加上信号与噪声过程均为高斯的限制条件，利用贝叶斯方法求解．与此同时，我们证明线性最小误差方差方法(实际上它是伪贝叶斯方法)得到的算法与贝叶斯估计算法相同．

运用使样本变稠密的方法，以及根据维纳-何甫方程和直

接应用变分法得到了连续时间的估计算法. 最后，获得经典的维纳滤波结果，并对估计算法的稳定性和渐近性态作了讨论.

第八章　最优线性滤波的推广　在这一章里，对第七章讨论的最优线性滤波作了若干推广. 第七章的基本假设是装置噪声与量测噪声均为白的. 对于有色装置噪声，采用附加状态矢量的方法就可解决，没有讨论的必要；如果观测（或量测）噪声为非白的，要确定最优滤波却是相当困难的. 这里，我们给出了处理各类有色量测噪声问题的一些方法.

第七章的序贯滤波方法，没有试图用稍后出现的观测去修改前面的估计量；但在许多问题中却希望这样做，我们称为平滑，在第八章中，我们介绍新息方法，以导出固定区间、固定滞后和固定点平滑的各种平滑的算法.

其次，我们讨论由于不适当地选取验前方差参数，以及不适当地选取消息模型与观测模型所造成的影响，并且确定出滤波与平滑的误差与灵敏度分析的算法. 最后，以发散性的讨论来结束本章. 发散现象是指状态的估计值与状态之差随着时间增大，这是由于依据不适当的验前统计量与不适当的消息、观测模型所造成的.

第九章　非线性估计　作为本书的最后一部分内容，对于非线性估计问题作一清晰的讨论. 重新研究第四章提出的随机微分方程概念，并对被估计状态矢量以观测为条件的情形导出了 Fokker-Planck 方程. 这个偏微分方程的近似解法就成为计算上可行的连续时间系统非线性滤波的算法. 其次我们证明，贝叶斯（或极大验后）离散估计等价于去做最小二乘曲线拟合. 我们得到了可用最优化方法求解的代价函数. 从而使我们获得非线性离散时间极大验后滤波的近似估计算法.

将非线性消息与观测模型次线性化，并应用线性滤波理论，我们获得了推广的连续时间与离散时间的非线性滤波算法．根据高斯随机变量的条件期望定理，得到了伪贝叶斯离散非线性滤波算法．然后，举例说明各种非线性滤波算法之间的相似性和差别．最后介绍了贝叶斯平滑算法和预测算法．

附录　给出两个附录作为补充内容．本书各处大量地用到矩阵求逆引理，在附录A中，证明了这个十分有用的引理．

本书某些章节需要最优化理论的概念．附录B向不熟悉这个领域的读者，概括地介绍最优化理论的有关内容．

参考文献　本书介绍了估计理论及其在通讯与控制中的应用的主要内容，并作了深入的阐述．在参考文献中，还会有更多的内容，并且已经介绍了更多的内容．文献目录按字母次序列出了这个领域中的许多有关的重要著作．文献名称后面标有本书相应的章节，以指导读者适合自己的兴趣范围去选取有关文章．

第二章 概率与随机变量

2.1 引 言

在任一通讯或控制系统中，它所处理信号的某些特征应是事先不知道的；否则，就不会有真正的信息传输(我们将信息定义为作出判定的有用资料)．本章将回顾概率论的部分内容，它们对于研究通讯与控制学科中的估计与判定理论是十分有用的[1]．

2.2 概 率

概率论是研究数学上可描述事件(或实验)的平均性态的．这种研究的基础是，当试验(或观测)的次数增加时，确信这些平均数接近于定值．例如，当掷钱币充分多次时，我们就相信掷出正面的次数差不多是总次数的一半．

为了清楚地描述概率的理论，有必要对含有机遇的实验给予确切的描述．我们把一个实验的结局称为结果．一个给定实验的所有可能结果的总体(或集合)称为样本空间 Ω．样本空间的子集称为事件．事件可以是简单的(不可分解的)或复合的(可分解的)．例如，掷骰子出现偶数点就是一个复合事件，因为它可以分解为三个简单事件：出现两点、出现四点、出现六点．

在一个实验中，事件已经发生的次数除以试验总次数称为事件发生的相对频率．当试验的次数趋于无限时，我们将

1) 对于概率与随机过程的较详尽论述，读者可参看列在参考文献中的专著．

相对频率称为事件发生的概率. 因而, 事件 $x=\alpha$ 发生的概率记为 $P_x(\alpha)$[1],

$$P_x(\alpha) \triangleq \lim_{N\to\infty} \frac{n_x(\alpha)}{N} \tag{2.2-1}$$

其中 $n_x(\alpha)$ 是事件 $x=\alpha$ 发生的次数, N 为试验的总次数. 函数 x 对应事件, α 对应事件的值. 为论述简单起见, 我们假定事件 α_i 是不相交的, 就是说没有两个事件 α_i 与 α_j 能够同时发生.

若有 M 个不相交事件 $\alpha_1, \alpha_2, \cdots, \alpha_M$ 产生样本空间, 则

$$N = \sum_{i=1}^{M} n_x(\alpha_i)$$

因此

$$1 = \frac{N}{N} = \lim_{N\to\infty} \frac{\sum_{i=1}^{M} n_x(\alpha_i)}{N}$$

$$= \sum_{i=1}^{M} \lim_{N\to\infty} \frac{n_x(\alpha_i)}{N} = \sum_{i=1}^{M} P_x(\alpha_i) \tag{2.2-2}$$

换言之, 产生样本空间的一组不相交事件的概率之和为1. 应该注意, 任何事件的概率总是非负的.

常常有两个实验同时进行, 并且我们希望考虑在第一个实验中发生事件 $x=\alpha$ 与第二个实验中发生事件 $y=\beta$ 的概率. 由(2.2-1)式易知

$$P_x(\alpha) = \lim_{N\to\infty} \frac{n_x(\alpha)}{N} \quad P_y(\beta) = \lim_{N\to\infty} \frac{n_y(\beta)}{N} \tag{2.2-3}$$

但是, 这两个概率并没有完全描述了实验的可能结果, 因为我们也可以考虑联合事件 $x=\alpha$ 与 $y=\beta$ 的概率. 如果我们假定这个联合事件发生了 $n_{x,y}(\alpha, \beta)$ 次, 则 $x=\alpha$ 与 $y=\beta$ 的概率[2]是

1) 通常, 一个事件的概率表示为 $P(A)$, 但这可能引起混淆, 因为我们希望清楚地区别函数与函数的值.

2) 当提到 $x=\alpha$ 的概率时, 意味着事件 $x=\alpha$ 发生的概率.