稳健信号处理概论

罗永光　王海云　编

国防科技大學出版社

内容简介

本书的目的是对稳健信号处理这一领域作一简要而系统的介绍，讨论在不确定噪声环境或不确定的谱特性条件下使信号的处理获得稳健特性的理论、方法及应用。回顾了稳健统计学发展的历史，描述了信号处理的各种稳健性概念，讨论了稳健估计和稳健检测的基本原理和方法，简述了稳健滤波、稳健量化、稳健谱估计、稳健阵处理等当前正处活跃的若干研究课题。

本书力求高的可读性，避免过多的数学推导，尽量结合实例着重阐述各种稳健处理方法的思路和物理含义。

本书适合于具有初步概率统计知识的读者作为学习稳健统计理论和研究稳健信号处理的导引性读物，适合于通信、控制、雷达、气象、地质、地震、经济管理等等多种领域与信号处理有关的工程技术人员作为参考，适合于高等院校师生作为教学参考书。

稳健信号处理概论

罗永光　王海云　编著

责任编辑　潘　生

封面设计　侯　云

*

国防科技大学出版社出版

国防科技大学印刷厂印装

湖南省新华书店发行

*

开本 787×1092 1/32　印张 8.44　字数 191 千字

1987年11月第一版　1987年11月第一次印刷　印数：0 001—3 000册

统一书号：15415・029

ISBN 7-81024-031-7

TN・2　定价：1.75元

序

数理统计学中稳健性概念的萌芽，可以从高斯提出正态律和最小二乘法的时候算起，但稳健统计的迅速发展并成长为数理统计学的一个分支，这还只是六十年代以来的事。

在应用统计方法解决实际问题时，常常受到两个方面的“困扰”：一是观测数据由于各种原因而存在异常值；二是数学模型与实际情况之间存在差异。当遇到上述情况时，经典统计方法中的统计量以及统计推断的结果往往会发生显著的变异，甚至导致不合理的或者错误的结果，许多最优的统计方法就不再是最优的了。总之，它们不能经受实际情况的种种变化，在这个意义上，我们说它们不够“强壮”和“坚韧”，也即缺乏“稳健性”。

于是，人们就试图寻求一类统计方法，使之具有下述特点：

(1) 在假定模型正确时，它具有良好的性质，是接近最优的；

(2) 在数据与模型差异较小时，它的性能变异也较小；

(3) 在实际情况偏离模型甚大时，它的性能也不会变得很差或导致破坏性的后果。

具有上述特征的统计方法，我们就称它为稳健的统计方法。

本书的书名是《稳健信号处理概论》。这是把稳健统计的思想和方法应用于信号处理领域，产生的信号稳健处理的应用理论。作者综合了稳健信号处理的一系列最新文献，结合作者个人在信号处理方面的实际工作，写出了这本书，它是一本导

引性的工程应用技术的著作。对于那些迫切于应用稳健统计理论来解决信号处理的科技工作者、大学生、研究生，无疑是一本很好的入门书。

作者学风严谨，立意明确，文笔流畅，我很高兴把它介绍给工程技术界的同志们，希望能对信号处理领域的深入研究有所裨益。

汪 浩

一九八七年二月于长沙

前　言

稳健统计学经历了长期的酝酿之后，从本世纪六十年代起进入了系统研究的阶段，将稳健统计学的理论和方法应用于信号处理的各个方面的研究已经全面铺开，稳健信号处理已经形成了一个广阔的领域。

稳健信号处理是介于信号处理的参量法与非参量法之间的一种新型处理方法。经典的最佳信号处理理论属于参量法，它基于对干扰和噪声的分布函数作理想化假定，因而在实际工程上如果干扰模型偏离理想情况，系统性能就会变差。非参量法由于没有利用或没有充分利用干扰的统计信息，因而存在保守性，没有充分发挥系统的潜力。稳健信号处理在解决这一矛盾中起着重要作用，是信号处理发展的一个重要分支，在工程实践上和理论上都具有很好的前景。

目前我国学术界和工程界虽已开始关心稳健统计学及其应用的课题，但此课题对大多数人来说还是较陌生的。特别是对于只愿对此领域作一般了解或虽想钻研但尚未入门的读者，很多稳健统计的文献其数学起点都嫌太高。他们希望有一本程度适中的较为系统的参考材料以作导引。

我们综合了稳健信号处理这一领域一系列最新文献，写出这本导引性的读物，目的是想对这一领域作简要而系统的介绍，希望能起到为浏览者引路、为进取者铺路的效果。书中揉合了不少作者自己对有关内容的理解，也许不尽恰当。如有谬误之处，恳请读者给予指正。

本书第一、二、三、五章由罗永光执笔，第四、六章由王

海云执笔，二人共同讨论定稿。

本书的写作与出版得到顾德仁教授与汪浩教授的帮助；陆仲良教授对本书的出版给予了支持。作者向他们表示深切的感谢。

目　录

第一章 绪 论

§1.1 稳健信号处理的含义

信号处理是从干扰或噪声中提取有用信号并进一步从中提取出人们所需要的各种信息的过程。

待处理信号来源于各种物理观测，而各种观测总是包含着若干不确定因素。比如说接收电报信号，虽然我们可以肯定这种信号必然是一些矩形脉冲，但由于通信信道中存在的干扰和噪声是不确定的，我们却不知道接收到的变了样的矩形脉冲究竟是信号还是干扰。又比如说天文信号，且不说观测中必然有其他的干扰混入，就是被观测天体所能给出的信号究竟是什么样子也是未 可 确 知 的。为了对付这些不确定性，人们常采用的办法是增加观测次数，对大量的观测结果作统计处理，由此产生了信号处理的统计推断理论。但是统计推断仅仅是部分地建立在观测的基础上的，另一个同样重要的基础乃在于对考虑情况的先验的假定。即使在最简单的情况下也存在着对于随机性、独立性、噪声背景的概率分布模型、某些未知参量的先验分布等的或明或暗的假定。这些假定往往是为了数学上的方便而引入的，或者是对实际情况所作的简化，或者只是对朦胧的知识作的粗略的描绘。这是一些理想化的假定，是信号处理的背景条件的理想化模型。

经典的信号处理方法就是基于对干扰或噪声的概率分布函数作理想化假定，在假定的理想噪声模型下设计对某种性能准则为最佳的信号处理器。当干扰或噪声的真实概率分布与理想

化的概率分布有差别时，处理效果是否还是最佳这一点，经典的处理方法是不大管的，或者认为只要真实分布与理想分布差别小，就能基本保持该统计处理过程的最佳性。其实这种直觉是成问题的。首先，两个概率分布函数之间的差别怎样就算是小，怎样就算是大，这问题在经典处理中不清楚。其次，模型或对分布的假定如何影响统计处理的结果，这点在经典处理中很少考虑。实践证明，某些经典的最佳处理方法，尽管是按理想化概率模型设计的，但当实际分布特性不同于理想分布时，处理结果也可以变化不大；而另一些经典的最佳处理方法，尽管我们直觉上并没有发现实际分布与理想分布之间有太大的差别，实际处理效果却会变得很差。这就是说，经典的信号最佳处理方法并不总是稳妥可靠的。

有一种信号处理方法，只对概率分布作很少的一般性假定，而不限制概率分布的具体形式，这种方法叫作非参数方法（相应地称过分地依赖于理想化概率模型假定的经典最佳处理方法为参数方法）。比如样本均值和样本中位数就分别是总体均值和总体中位数的非参数估计。有些非参数方法（如样本中位数）是比较稳定的，但也有些非参数方法（如样本均值）对于异常的观测值极为敏感，很不稳定。同时，也因为这种方法对具体的概率分布情况考虑很少，就没有充分地利用实际工作中可以得到的有关干扰或噪声的统计观测知识，在实际分布与理想化模型差别不大时，就不如参数方法那样有效，还存在着相当的保守性和盲目性。

信号的稳健处理方法，应该克服参数方法和非参数方法的缺点而综合二者的长处。就是说，这种方法要充分地利用可以得到的统计知识，利用这些知识以建立适当的概率分布模型，此模型应考虑到实际的概率分布可能出现的偏离，然后在此模

型的基础上设计出信号的最佳（或次佳）处理方式。概括起来，稳健的信号处理应具有如下的三条性质：

1. 在假定模型符合实际情况时，应具有最佳（或次佳）的处理性能；

2. 在实际情况与假定模型差别较小时，处理性能变化也较小；

3. 在实际情况偏离假定模型较远时，处理性能不能变得很差或造成破坏性的影响。

注意，这里只提假定模型而不指明什么模型，因为在信号处理中除了概率分布形式的假定外，还可能有其他的假定，如谱分布假定等。

总之，信号的稳健处理方法，既应该是稳定的，也应该具有尽可能最佳的处理效果。“稳健”这个词的英文为Robust，含义是强壮，健康、坚韧、能经受逆境的考验。也有人把Robustness（稳健性）译作顽健性、坚韧性。但要注意，稳健概念首先是求稳，在稳中求健（最佳）。

信号处理过程的稳健性与其可用性是紧密相关的。为了确定某一给定信号处理过程的适用范围有多宽，可以观察它究竟对哪一些假定方面的偏差表现稳健。而如果把一处理过程设计得稳健，就等于说尽管在我们获得的信息中可能存在不确定因素，我们设计的处理过程也能保持可靠。

在信号处理中，还有一些与稳健方法相近的方法，比如自由分布（distribution-free）方法、自适应(adapted)方法、抗变性（resistant)方法等。自由分布方法要求处理效果对于所有可能的基准分布都相同；自适应方法是使处理过程能自动地适应于基准模型；抗变性是指处理结果对基准样本的微小变化(全体的微小变化或少数数值的大变化）不敏感。稳健方法与参数

方法的接近程度远甚于与非参数方法及自由分布方法的接近程度，因为稳健方法必然是按某种模型进行的，尽管它并不囿于此种模型而要考虑模型偏差的影响。有人（如W.J.J.Rey）把自适应方法纳入稳健方法[21]，而有人（如P.J.Huber）则主张将稳健方法与自适应方法区分开，认为前者更多地强调安全性（稳定性）而不是有效性[8]。抗变性与稳健性在概念上虽稍有区别，但这两个概念对一切实际目的来说都是同义的[8]。

§1.2 稳健信号处理的发展简况

信号的稳健处理正在成为一门新兴的学科，它是将数理统计学的新的理论分支——稳健统计学（Robust statistics）应用于信号的分析处理的一门理论与技术相结合的学问。稳健信号处理的发展史，与稳健统计学的发展史是分不开的。

让我们先来看看稳健统计方法的思想发展史。

稳健统计方法的思想发展史是与经典统计方法的发展史交织在一起的。

1809年，高斯(Gauss)导出了测量误差的正态分布。对于该分布，观测值的样本均值是分布中心在均方意义下的有效估值。高斯利用样本均值的最主要理由是计算简单，因而这种方法在后来得到了广泛的实际应用。但是勒让德（Legendre）在1805年发表第一篇关于最小二乘法的文章时就指出，在利用这些方法（特别是样本均值）之前，应该仔细地查看一下样本，抛弃那些是、或者似乎是反常的远离值。1818年，贝塞尔（Bessel）研究了一系列的大样本后得出结论说，大的测量误差比起可能服从正态分布的误差值会更经常地碰到，然而人们都忽略了这一事实。

1818年，拉普拉斯（Laplace）对于测量误差服从任意对

称分布的情况，提出用最小模方法来解单参数（斜率）的线性回归问题。他找到了构造斜率参数估值的分布，并且证明了估值的渐近正态性，还得到了此时的渐近方差小于最小二乘法估值的渐近方差的条件。

佩尔斯(Peirce)和乔维奈特（Chauvenet）分别于1852年和1863年提出了如何找出和剔除异常测量值的方法。剔除异常数据方法的主要缺点在于，它们是相当复杂的，并且当存在大量的异常值（5～20%）以及当分布有沉重的拖尾时，这种方法事实上是无用的。

1886年，牛柯姆（Newcomb）根据天文观测值的分析得出结论：实际测量值的分布比起正态分布来有更重的拖尾。为了逼近这些分布，他提出利用某些正态分布的混合形式，这些正态分布的方差不同而均值相同，而且应用有均匀先验分布的贝叶斯估值作为此混合分布的均值的估值。此后，牛柯姆又研究了一种很简单的线性加权估值方法，他所提出的方法原则上很接近于构成稳健估值的现代方法。

1886年，艾吉沃斯（Edgeworth）发现，当测量误差相关时，样本中位数可以给出比样本均值更好的估值。

1889年，高尔顿（Galton）提出由两个样本百分位点的线性加权形式作为均值和标准偏差的简单估值。谢帕德（Sheppard)证明了当测量值服从正态分布时，高尔顿估值具有联合渐近正态性，并且把这种估值推广到任意数目的样本百分位点的线性加权情况（类似后来的L估计），还试图使这种估值具有最佳的性质。

1912年，庞加莱（Poincare）提出了截尾均值，即在构造估值时把最边缘部分的样本值去掉。有趣的是，类似的估值老早就已在法国用来计算粮食年平均产量了。

1920年，丹尼尔（Daniel）提出了采用L估值（次序统计量的线性加权和）来估计总体的均值和方差，证明了L估值的渐近正态性，还作出了使L估值的渐近方差最小的最佳权函数。到1955年江(Jung)的文章发表为止，这篇富有内容的文章一直被人们忽略了三十五年之久。只是由于郭德文(Godwin)、洛伊德（Lloyd）的文章的发表，才使人们对L估值重新发生兴趣。

1931年，皮尔逊（Pearson）以及近来的其他作者证明了经典的假设检验方法对其最佳性条件的微小破坏是高度敏感的。

在本世纪四十至六十年代期间，以秩统计量为基础的非参数方法取得了飞快的发展。1963年，霍吉斯（Hodges）与列曼（Lehmann）提出了用秩方法以构造稳健估值。

在稳健统计方法的发展过程中，普林斯顿大学的J·W·图基（J·W·Tukey）和他的统计学讨论班在五十年代末所进行的工作起了巨大的作用。他们的工作标志着稳健统计学的酝酿阶段的结束，预示着稳健统计学的系统性研究阶段即将开始。

如上所述，早在提出正态分布律和最小二乘法的时候就有了稳健性思想的萌芽，到本世纪五十年代为止，稳健统计学已经历了一个半世纪的酝酿阶段。不过，稳健统计学得以迅速发展并成为新兴的数理统计学分支，则还是近二三十年以来的事，其发展是得力于非线性数学研究的进展和计算机迭代算法的进步。

稳健性(robustness)这个术语最早是G·E·P·薄克斯(Box)在1953年引入的。他的目的是想描述统计中一种当时尚不清晰的、类似于稳定性一类的概念。而冯米赛斯（Von Mises）（1942）和普洛合洛夫（Prokhorov）（1956）讨论了估值的

渐近理论和有限容量的实际情况之间的关系，为统计稳健性理论的发展提供了简明的轮廓线[21]。

六十年代以来，稳健统计学的研究出现了热烈的局面。稳健统计学的系统性研究一般公认为始于1964年，这一年，P·J·胡倍尔（Huber）发表了“位置参数的稳健估计”[6]的开创性论文。他的这篇文章第一次正面回答了应该如何设计稳健统计过程的问题。第二年即1965年，他又发表了一篇叫作“概率比检验的稳健形式”的论文[7]，从而开始了稳健性理论在信号检测中的应用研究。

F·R·汉培尔（Hampel）1968年的学位论文联系冯米赛斯的工作，引入了影响曲线（函数）这一重要概念。按P·J·胡倍尔的说法，这也许是稳健统计学中一个最有用、最富启发性的工作了。

从七十年代起，稳健信号检测与估计的研究全面开展起来了，研究的人多了，研究的范围也广了。国际上在有关的杂志及会议文集中，研究稳健统计与稳健信号处理的文章源源不断。

1981年，P·J·胡倍尔出版了第一本系统地论述稳健统计学的专著[8]。1983年，W·J·J·瑞（Rey）在他自己1978年的专著的基础上又出版了一本谈论稳健和次稳健方法的书[21]。

如果说七十年代末以前有关的研究主要是针对概率分布假定方面的稳健性，那么八十年代以来，针对其他假定的稳健性的研究也活跃起来了。S·韦度（Verdú）和H·V·普尔（Poor）等从对策论的角度将极小极大稳健性的思想一般化，将这种方法推广应用到关于谱分布等的不确定性的匹配滤波以及其他一些信号处理问题[22]。有关稳健维纳滤波、稳健卡尔曼滤波等课题的研究也正在不断地深入。

随着稳健信号处理理论研究的深入，其应用研究也不断地在深入。这里可以举出一些应用稳健方法的实际例子。例如：R·D·马廷（Martin）和C·J·马斯瑞利兹（Masreliez）用稳健方法对宇航装置的飞行轨道参数提供可靠的估计；Д·А·布拉斯拉夫斯基（Браславский）等用稳健方法构造处理过剩数据的可靠系统；L·F·费霍勒(Fehler)等利用截尾均值来提高导航系统的可靠性；W·S·克列维兰德（Cleveland）等应用稳健方法来分析大城市的空气污染数据；M·麦瑞尔（Merrill）等在电力系统的研究中应用M估计和各种近似的误差分析方法；W·S·阿吉（Agee）用稳健估计解决弹道数据的平滑问题，等等。

§1.3 本书概况

本书的目的是简要而系统地介绍稳健信号处理的概念和方法。全书分六章：第一章，绪论。第二章、讨论观测信号的概率模型、污染分布、分布拖尾的影响、效率稳健性、分布函数的邻域及距离、信号处理器的定性稳健性、崩溃点、影响函数以及极小极大稳健性的概念。第三章介绍四种稳健估计方法：M估计、R估计、L估计以及递推估计（特别是其中的随机逼近估计）。其中，前三种是基本的稳健估计方法，尤其是M估计较成熟，也用得较多。而随机逼近（SA）估计则由于可以递推实现，便于实时处理，近年来很为学术界所重视。第四章讨论稳健信号检测原理。对具有极小极大性的下列一类稳健检测器的工作原理和设计方法作了较详细的叙述：概率比稳健检测；采用局部最大势检验的非线性相关检测；分别利用M估计、SA估计作为检验统计量的稳健M检测和稳健SA检测；稳健M序贯检测和稳健SA序贯检测。此外，对随机信号及带通

信号的稳健检测问题也进行了讨论。信号处理是一个十分宽广的领域，为使读者对信号稳健处理问题的普遍性有个初步的了解，我们在第五、六两章中，对不确定概率分布下的其他课题：稳健回归、稳健卡尔曼滤波与信号建模及系统辨识、稳健谱估计、稳健量化，以及一类非概率分布的不确定性、即不确定谱分布下的稳健处理课题：稳健维纳及匹配滤波、稳健均衡、稳健阵处理等领域的概况也作了简明的介绍。

本书适合于有一点概率统计知识的读者作为了解稳健统计理论和研究稳健信号处理的导引性读物，适合于雷达、通信、控制、气象、地震和经济管理等领域与信息处理有关的大学师生和工程技术人员作参考。读完本书，将可对稳健信号处理的基本概念和实现方法以及其广泛的应用有个初步的较为系统的了解。