

设:

$\gamma(\frac{2}{3})\dots$

, 柯

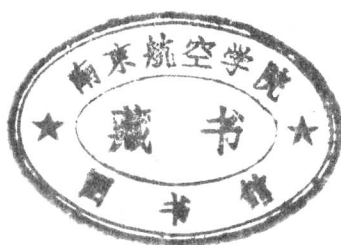
卡尔曼滤波

及其在惯性导航中的应用

卡尔曼滤波 及其在惯性导航中的应用

组成部分。
息，并且
密各种

俞 济 祥 主 编



航空专业教材编审室

512021

内 容 简 介

介绍了卡尔曼滤波的基本理论及其实际应用问题,并以惯性导航为例,介绍了滤波应用方法。第一篇滤波基本原理,大致分为三部分内容:滤波基本原理;滤波在实际应用中的某些问题和在随机最优控制中的应用和在非线性系统中的推广应用。第二篇滤波在惯性导航系统中的应用,导初始对准和组合导航应用卡尔曼滤波的具体方法。按照航空工业部航空高等院校教材统编规划的要求编写的,是《惯性导航与仪表》专业高年级的主材及该专业研究生的必修教材,也可作为导航和其他技术领域的工程技术人员工作的参考书。

主审稿人 申功勋

责任编辑 肖龙跃

卡尔曼滤波及其在惯性导航中的应用

俞济祥 主编

*

航空专业教材编审室

西北工业大学印刷厂印装 内部发行

*

787×1092 1/16 印张 15.125 字数 374 千字

1984年11月第一版 1984年11月第一次印刷 印数 0001—1000 册

统一书号: 36235 j 定价 1.56 元

前 言

卡尔曼滤波是六十年代初发展起来的一种估计方法，是现代控制理论的重要组成部分。由于它采用状态空间和递推计算等方法，所以，能估计平稳和非平稳随机过程的信息，并且有数据贮存量少以及易于用计算机实现等优点。利用卡尔曼滤波可以得到随机动态系统各种信息(状态)的线性最小方差估计，而且，与最优控制相结合，可以实现随机动态系统的最优控制。目前，国外在空间技术和工业生产等领域广泛应用了卡尔曼滤波。惯性导航由于其本身的特点：可用数学方程确切描述系统的动态特性，动态特性变化较缓慢，以及可用实验方法取得噪声统计知识等，使之成为卡尔曼滤波最早得以应用并得到成功的领域之一。

本书以惯性导航为对象介绍卡尔曼滤波。由于滤波在实际应用中有各种具体问题和处理技巧，因此，书中除了阐述滤波基本理论外，还着重介绍了应用问题。书中内容参考了有关卡尔曼滤波及其在惯导方面应用的书籍和文献，尽可能反映了目前这一领域的主要研究成果。全书共分两篇，第一篇阐述滤波基本理论和实际应用中的有关问题和处理方法，诸如有色噪声的卡尔曼滤波、序贯处理、信息滤波方程、发散问题、平方根滤波、各种次优滤波方法，以及在设计滤波器时必须掌握的误差分析方法。对有关理论的证明，尽可能采用简明的方法，并阐述了证明的思路。第二篇以惯导系统的初始对准和组合导航为例，介绍实际应用卡尔曼滤波的步骤和方法。其中有关系统可检测性的判别与分析，系统离散化中转移矩阵和系统噪声方差阵的计算方法，噪声模型的确定和简化，以及量测值平均等方法，对其他领域的滤波应用也有参考价值。为了概要说明每章的主要内容，在每章末尾都有小结，并附有习题。

附录共四个：附录A介绍连续系统离散化问题及计算方法，附录B和附录C是为第五章滤波稳定性提供的参考内容，附录D是为第八章中平方根滤波提供的参考内容，也提供了矩阵秩的计算方法。

本书是在西北工业大学和南京航空学院两校原有讲义的基础上编写的。其中第十一章由南京航空学院汪叔华执笔，其余部分由西北工业大学俞济祥执笔，全书俞济祥主编。本书承北京航空学院申功勋同志详细审阅，并提出了宝贵的意见，西北工业大学赵维华同志为本书绘制了全部插图，在此谨致深切谢意。

书中谬误和不妥之处，请读者批评指正。

编 者
一九八三、六

目 录

第一篇 卡尔曼滤波基本原理

第一章 绪论	1
§ 1.1 最优滤波介绍.....	1
§ 1.2 卡尔曼滤波在惯性导航系统中的应用.....	3
第二章 几种最优估计	5
§ 2.1 最小二乘估计.....	5
§ 2.2 最小方差估计.....	10
§ 2.3 线性最小方差估计.....	15
第三章 离散系统卡尔曼滤波方程	20
§ 3.1 离散系统卡尔曼滤波基本方程的直观推导.....	20
§ 3.2 正交投影原理.....	31
§ 3.3 离散系统卡尔曼滤波基本方程的正交投影推导.....	34
§ 3.4 离散系统卡尔曼滤波基本方程的几点说明.....	35
§ 3.5 白噪声情况下离散系统卡尔曼滤波一般方程.....	37
第四章 连续系统卡尔曼滤波方程	42
§ 4.1 连续系统的描述.....	42
§ 4.2 连续系统卡尔曼滤波基本方程的推导.....	44
§ 4.3 连续系统卡尔曼滤波基本方程的几点说明.....	46
§ 4.4 连续系统卡尔曼滤波一般方程.....	48
第五章 卡尔曼滤波稳定性介绍	52
§ 5.1 问题的提出.....	52
§ 5.2 滤波稳定的几种判别条件.....	52
§ 5.3 适用于惯性导航系统的滤波稳定判别条件.....	59
第六章 随机线性系统的最优控制	63
§ 6.1 概述.....	63
§ 6.2 离散系统的分离定理.....	64

§ 6.3	连续系统的分离定理	68
§ 6.4	估计直接反馈控制	69
第七章	有色噪声条件下的卡尔曼滤波	74
§ 7.1	有色噪声	74
§ 7.2	有色噪声条件下的卡尔曼滤波	79
第八章	在滤波计算中有关滤波方程的一些考虑	85
§ 8.1	序贯处理	85
§ 8.2	信息滤波方程	86
§ 8.3	发散	87
§ 8.4	平方根滤波	91
第九章	卡尔曼滤波误差分析与次优卡尔曼滤波	98
§ 9.1	卡尔曼滤波误差分析一般方法	98
§ 9.2	模型正确时的误差分析方法——误差范围	105
§ 9.3	降阶次优滤波方法	107
§ 9.4	其它次优滤波方法	112
第十章	最优线性平滑	117
§ 10.1	概述	117
§ 10.2	固定点平滑	117
§ 10.3	固定滞后平滑	120
§ 10.4	固定区平滑	123
第十一章	非线性系统的卡尔曼滤波	127
§ 11.1	引言	127
§ 11.2	按标称状态线性化的卡尔曼滤波方程	128
§ 11.3	按最优状态估计线性化的卡尔曼滤波方程——广义卡尔曼滤波方程	133
§ 11.4	举例	139

第二篇 卡尔曼滤波在惯性导航系统中的应用

第十二章	惯性导航系统误差方程	141
§ 12.1	矢量差误差方程	142
§ 12.2	分量差误差方程	147
§ 12.3	关于误差方程的几点说明	150

.....68		
.....69	十三章 卡尔曼滤波在惯性导航系统中应用的几种结构形式	152
.....74	§ 13.1 直接法和间接法	152
.....74	§ 13.2 输出校正和反馈校正	155
.....79	§ 13.3 惯导系统采用间接法滤波必须注意的问题——卡尔曼滤波用于惯导系统自阻尼方面的限制	157
.....85	十四章 卡尔曼滤波在惯性导航系统初始对准中的应用	160
.....85	§ 14.1 概述	160
.....86	§ 14.2 静止基座状态方程和量测方程的列写	161
.....87	§ 14.3 摆动基座状态方程和量测方程的列写	167
.....91	§ 14.4 离散化的问题	168
.....98	§ 14.5 关于将估计用作校正的具体实施问题	170
.....98	§ 14.6 举例	172
.. 105	§ 14.7 次优滤波器	175
.. 107	十五章 卡尔曼滤波在组合导航中的应用	178
.. 112	§ 15.1 各种导航系统简介	179
.. 117	§ 15.2 系统可检测性的判别	181
.. 117	§ 15.3 状态的选择和状态方程的建立	183
.. 117	§ 15.4 量测方程的建立	188
.. 120	§ 15.5 噪声模型的确定	193
.. 123	§ 15.6 反馈校正的实现	195
.. 127	§ 15.7 举例	197

附 录

127		
127		
128		
133	附录 A 连续型状态方程的离散化	207
139	一、定常系统状态方程的离散化	207
	二、时变系统状态方程的离散化	208
	三、定常系统转移矩阵 $\phi(T)$ 和噪声方差阵 $\bar{Q}_k$ 的计算方法	209
	附录 B 离散型线性系统的稳定性	212
41	一、稳定性的定义	212
42	二、稳定性的充要条件	213
47	附录 C 线性系统的可控性、可观性、可稳定性与可检测性	216
50	一、可控性	216

二、可观测性.....	219
三、定常系统的完全可稳定性.....	222
四、定常系统的完全可检测性.....	224
附录 D 矩阵的正交三角化	225
一、镜象变换.....	225
二、用镜象变换可使任何矢量 ξ 变换为与给定单位矢量 γ 同方向的矢量.....	227
三、矩阵的正交三角化.....	228
四、利用矩阵正交三角化计算矩阵秩的方法.....	230
参考文献	231

第一篇 卡尔曼滤波基本原理

第一章 绪 论

§ 1.1 最优滤波介绍

1. 最优滤波

从量测到的或接收到的各种带有干扰的信号中取出有用信号的方法或技术称为滤波，或称为信号滤波。收音机的频率调谐就是按照需要的电台频率将信号从无线电波信号中取出来的一种信号滤波。信号滤波普遍应用于通讯技术中，一般以电路的形式具体体现在设备中。这里要介绍的是根据统计特性的规律，将所需信号从含有干扰的量测中得出的一种信号滤波，称为统计滤波。确切地说，这种滤波是根据所需信号和量测值的统计特性，从量测值中得出某种统计意义上误差性能最小，亦即“最优”的信号估计，一般称为最优滤波或最优估计。根据不同的最优准则和估计方法，有各种不同的最优滤波。它有电路形式的滤波器（如维纳滤波，它采用电子放大器电路），但更一般的是用数字计算的方法来计算和估计出所需信号的数值，如果说有滤波器，指的就是用以计算数学方程和实现方程解算的工具（如数字计算机）。

2. 最优滤波的发展过程

早在1809年，高斯为了从已知量测数据中确定天体运行的轨道，就提出了一种最早的最优滤波方法，称为最小二乘法。虽然它的滤波性能较差，但因为它具有不需要信号的验前统计知识这个突出优点，所以至今尚在不少技术领域中使用。在惯性导航中，不论对元件测试的数据处理，还是初始对准或定期标定等工作中，有时仍然使用这种方法。

二十世纪四十年代，维纳为了火力控制上的需要而提出一种线性最小方差滤波方法，称为维纳滤波。虽然这种方法充分利用了信号和量测值的统计特性，但因它仅适用于平稳随机信号，而且，最优滤波器在工程上很难完全实现，因此，这种滤波理论的实用价值受到一定的限制。

1960年出现的卡尔曼滤波也是一种线性最小方差滤波方法。它不但考虑了信号与量测值的基本统计特性（一、二阶统计特性），而且，由于采用了状态空间的概念，用状态方程描述系统，信号作为状态，所以，它既能估计平稳的标量信号随机过程，又能够估计非平稳的矢量信号随机过程。卡尔曼滤波另一个特点是采用递推计算的方法。它不要求贮存过去的量测值，只要根据当时的量测值和前一时刻的估计，按照一套递推公式，利用数字计算机这个

现代计算工具，就可以实时地计算出所需信号的估计。因此，卡尔曼滤波又称为递推线性最小方差估计。图 1.1 概略的表示卡尔曼滤波的工作原理。可以看出，卡尔曼滤波的对象是用状态方程描述的随机线性系统。它按照估计误差的方差最小的准则，从被量测噪声“污染”

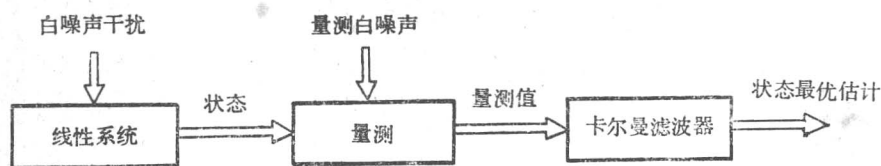


图 1.1 卡尔曼滤波原理示意图

了的量测值中实时估计出系统的各个状态值。它是根据概率论、数理统计和线性代数等数学理论发展起来的一门新兴技术，是现代控制工程中的一个主要方面。由于它具有能够估计非平稳的随机矢量过程，可以在计算机上实时计算，以及贮存容量少等优点，因此，一经提出，就很快在空间技术中（例如飞行器导航系统、导弹制导、测定人造卫星轨道和姿态以及火力控制等方面）得到了应用。七十年代以来，卡尔曼滤波又在动力系统、工业生产的最优控制和气象预报等方面得到了推广应用。

3. 最优滤波问题的分类

按量测值与被估计值的时间关系，最优滤波问题可分为三类，即最优滤波（或最优估计）、最优预测和最优平滑。

设 $\mathbf{x}(t)$ 为被估计的随机矢量， $\mathbf{Z}(t)$ 为量测随机矢量*，量测方程为：

$$\mathbf{Z}(t) = \mathbf{H}(\mathbf{x}(t), \mathbf{V}(t), t)$$

式中 $\mathbf{V}(t)$ 为量测噪声矢量， $\mathbf{H}(\cdot)$ 是 $\mathbf{x}(t)$ 、 $\mathbf{V}(t)$ 和 t 的已知矢量函数。最优滤波就是利用 t_0 到 t 过程中的量测值 $\mathbf{Z}(\tau)$ ($t_0 < \tau \leq t$)，通过一个滤波器去估计 t' 时刻的 $\mathbf{x}$ 值，使之在一定统计意义下估计 $\hat{\mathbf{x}}(t'|t)$ 为最优。如果 $t' = t$ ，则这种滤波称为最优滤波或最优估计， $\hat{\mathbf{x}}(t'|t)$ 称为 $\mathbf{x}(t')$ 的最优滤波值或最优估计；如果 $t' > t$ ，则这种滤波称为最优预测；如果 $t' < t$ ，则这种滤波称为最优平滑。而最优滤波是这三类滤波的统称，有时亦表示其中的最优估计。本书主要讨论最优估计问题。

估计问题大致可以分为两类：状态估计和参数估计。状态和参数没有严格的定义加以区别，一般把随时间变化的信号或物理量称为状态，而参数表示信号和信号之间的关系，它一般不随时间而变化或只有缓慢的变化。例如，欧姆定律： $E(t) = RI(t)$ ，式中 $E(t)$ 和 $I(t)$ 分别是电压和电流， R 是电阻，前两者是状态，后者是参数。如果已经用状态方程描述了一个线性系统，则状态和参数就明确地区分出来了。状态估计常常是动态估计，而参数估计一般是静态估计。对随机量来讲，动态估计就是随机过程的估计，而静态估计则是随机变量的估计。书中讨论的内容仅限于状态估计。

* 本书中随机矢量指的是矢量形式的多维随机变量。且 $\mathbf{x}$ 为被估计量， $\mathbf{Z}$ 为量测量，在以后的讨论中不再说明。

§ 1.2 卡尔曼滤波在惯性导航系统中的应用

惯性导航系统（简称惯导系统）是近三十多年来发展起来的一种自主式导航设备。它能够实时输出位置和速度等导航参数，并且为运载器提供精确的姿态信息基准。目前，在航空、航海和航天领域里有极为广泛的应用。随着科学技术的发展，对惯导系统不断提出了更高的要求。例如惯导系统在使用前要进行初始对准。对准的时间一般需要十几分钟，而且要求系统处于比较静止的环境。为了满足航行或战术的需要，要求对准时间更短。有的还要求能在动荡的环境中（例如在舰船上）得到又快又准的对准效果。但是，过去对准回路常常满足不了这些要求。又如，惯导系统的导航参数误差一般随时间而积累。长时间工作的惯导系统往往不能满足导航要求。导航精度主要受惯性元件（陀螺和加速度计）精度的限制，设法研制高精度的惯性元件固然是一条途径，但这需要化费高昂的经济代价。最适用的方法是组合导航，即利用惯导系统和其他导航设备组合成组合导航系统，既可以发挥惯性导航能够连续输出导航参数和姿态信息的优点，又可以保证精度而满足要求，还能提高导航设备的可靠性。此外，从经济效益角度出发，目前亦提出了利用低精度的惯导系统和其他导航设备组合成中等精度组合导航系统的方案。但是，采用校正环节组合的组合导航系统，不能充分发挥各自系统的作用。以上这些问题，如果采用卡尔曼滤波技术，可以得到很好的效果。下面简要说明卡尔曼滤波是如何在组合导航中使用的。

设惯导系统和多卜勒导航设备组合，卡尔曼滤波器的作用是估计惯导系统的导航参数误差。因此，组合的方式是将惯导系统输出的地速 $V_I(t)$ 与多卜勒雷达输出的地速 $V_D(t)$ 相减，其差值就是滤波器的量测值。设惯导系统地速的误差为 $\Delta V(t)$ ，多卜勒雷达的地速误差为 $\Delta V_D(t)$ ，则量测值 $Z(t)$ 为：

$$\begin{aligned} Z(t) &= V_I(t) - V_D(t) = [V(t) + \Delta V(t)] - [V(t) + \Delta V_D(t)] \\ &= \Delta V(t) - \Delta V_D(t) \end{aligned} \quad (1.2.1)$$

式中 $V(t)$ 是运载器的真实地速。我们将惯导系统的一整套误差（微分）方程用状态方程的形式来描述惯导系统，包括速度误差 $\Delta V(t)$ 在内的导航参数误差作为状态，并设 $\Delta V_D(t)$ 是白噪声，则上式说明量测值 $Z(t)$ 能量测出系统状态 $\Delta V(t)$ ，而 $\Delta V_D(t)$ 就是量测误差。产生导航参数误差的原因主要是惯导元件等误差源的影响。在掌握了这些误差源和 $\Delta V_D(t)$ 的

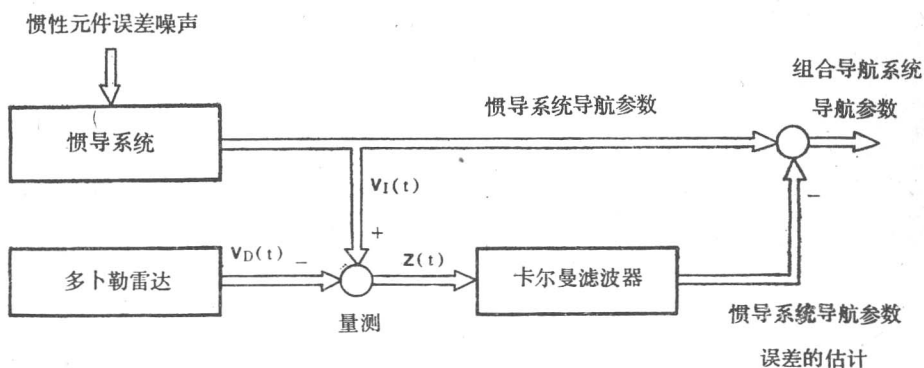


图 1.2.1 用卡尔曼滤波器组合惯导系统和其他导航设备的原理示意图

一、二阶统计特性后，卡尔曼滤波器就能根据这些统计特性的数值和系统状态方程，从量测值 $Z(t)$ 中实时估计出各个导航参数误差，用以修正惯导系统输出的导航参数。经过修正后的数值，就是组合导航系统输出的导航参数。从统计意义上讲，它要比惯导系统输出的导航参数精确得多。

从上面介绍的用卡尔曼滤波器来组合两种导航设备的简要原理来看，滤波必须有两个前提，一是滤波的对象能够较准确的由状态方程来描述，二是必须掌握系统噪声的一、二阶统计特性。而对惯导系统来讲，它的一整套一阶近似的误差微分方程能够较准确的描述惯导误差的规律。用它来描述卡尔曼滤波所要求的滤波对象是非常合适的。而且，滤波所要求的统计特性就是惯性元件误差和多卜勒雷达地速误差的统计特性。一旦这些产品定型后，统计特性可以从产品的实验结果中得出。可以说，惯导系统应用卡尔曼滤波的条件是十分有利的。所以，惯导系统是卡尔曼滤波最早应用并得到成功的一个领域。在七十年代，惯性导航又在大地测量技术等方面得到了成功的应用，卡尔曼滤波同样起到了不可低估的作用。

本章小结

最优滤波是信号滤波的一种，亦称最优估计。根据最优意义的不同，有各种不同的最优滤波。它估计的对象是随机线性系统。根据滤波目的的不同，它还可分为状态估计和参数估计两类。前者是动态估计，后者一般属于静态估计。卡尔曼滤波是最优滤波的一种，亦称为递推线性最小方差估计。卡尔曼滤波在惯导系统中得到了成功的应用。

第二章 几种最优估计

根据最优意义的不同,有各种不同的最优估计。在讨论卡尔曼滤波以前,本章先介绍几种常见的最优估计。它们是最小二乘估计、最小方差估计和线性最小方差估计。在一般情况下,它们仅适用于静态估计。因为它们的估计性能与卡尔曼滤波有相似之处,所以,了解这些估计方法,有助于讨论卡尔曼滤波,亦有助于理解卡尔曼滤波为什么是动态估计的原因。

§ 2.1 最小二乘估计

根据估计方法的简易程度以及量测数据处理次序的不同,最小二乘估计又可分为一般最小二乘估计、加权最小二乘估计和递推最小二乘估计。下面分别进行讨论。

2.1.1 一般最小二乘估计

在日常工作中,我们经常用“平均”的办法来简单处理数据,这就是一般最小二乘估计的方法。例如,对某一电压两次量测的结果,分别是 215V. 和 219V., 我们常认为正确的电压就是这两个量测值的平均值,即 217V., 这个数值就是一般最小二乘估计,常称为最小二乘估计。下面具体叙述。

设被估计量 $\mathbf{x}$ 是 n 维的确定值矢量, 并设对 $\mathbf{x}$ 进行了 k 次量测, 量测矢量 $\mathbf{Z}_i$ 与 $\mathbf{x}$ 有线性关系, 量测误差为 $\mathbf{V}_i$, ($i=1, 2, \dots, k$), 则量测方程分别为:

$$\begin{aligned}\mathbf{Z}_1 &= \mathbf{H}_1 \mathbf{x} + \mathbf{V}_1 \\ &\vdots \\ \mathbf{Z}_k &= \mathbf{H}_k \mathbf{x} + \mathbf{V}_k\end{aligned}\quad (2.1.1)$$

并设 $\mathbf{Z}_i$ 为 m_i 维, 则量测矩阵 $\mathbf{H}_i$ 是 $m_i \times n$ 矩阵, 将 k 个量测方程合并成一个, 即:

$$\mathbf{Z} = \mathbf{H} \mathbf{x} + \mathbf{V} \quad (2.1.2)$$

式中:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 \\ \vdots \\ \mathbf{Z}_k \end{bmatrix} \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_k \end{bmatrix} \quad \mathbf{V} = \begin{bmatrix} \mathbf{V}_1 \\ \vdots \\ \mathbf{V}_k \end{bmatrix}$$

则 $\mathbf{Z}$ 和 $\mathbf{V}$ 的维数等于 $\sum_{i=1}^k m_i = M$, $\mathbf{H}$ 为 $M \times n$ 矩阵。现在就根据合成的量测矢量 $\mathbf{Z}$ 对 $\mathbf{x}$ 作出估计 $\hat{\mathbf{x}}$ 。最小二乘估计的准则就是要求作出这样的估计: 使得各次量测值 $\mathbf{Z}_i$ 与估计 $\hat{\mathbf{x}}$ 对应的量测估计 $\mathbf{H}_i \hat{\mathbf{x}}$ 之差的平方和最小, 即:

$$J(\hat{\mathbf{x}}) = (\mathbf{Z} - \mathbf{H} \hat{\mathbf{x}})^T (\mathbf{Z} - \mathbf{H} \hat{\mathbf{x}}) \quad (2.1.3)$$

为最小。为求得估计 $\hat{\mathbf{x}}$, 先求 $J(\hat{\mathbf{x}})$ 的极小值, 则:

$$\frac{\partial J(\hat{\mathbf{x}})}{\partial \hat{\mathbf{x}}} = -2\mathbf{H}^T (\mathbf{Z} - \mathbf{H} \hat{\mathbf{x}}) = 0 \quad (2.1.4)$$

设 $(\mathbf{H}^T \mathbf{H})^{-1}$ 存在, 得到:

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{Z} \quad (2.1.5)$$

即按照上式计算得到的估计 $\hat{x}$ 满足(2.1.4)准则。 $(H^T H)^{-1}$ 存在即相当于系统是完全可以观测的。因为对确定值矢量 x ，可看作是转移矩阵 Φ 为 I 单位阵的系统状态。这样的系统如果是完全可以观测的，就相当于可观测阵 $M(k, 1)$ 是正定的。(参阅附录(c.2.1)式)，即：

$$M(k, 1) = \sum_{i=1}^k H_i^T H_i = H^T H > 0 \quad (2.1.6)$$

最小二乘估计有以下两种性质：

(1) 如果量测噪声 V 的均值为零矢量，即：

$$E\{V\} = 0$$

式中 $E\{\cdot\}$ 为均值符号，则估计误差 $\tilde{x}$ 为：

$$\begin{aligned} \tilde{x} &\triangleq x - \hat{x} = (H^T H)^{-1} H^T H x - (H^T H)^{-1} H^T Z \\ &= (H^T H)^{-1} H^T (H x - Z) \\ &= - (H^T H)^{-1} H^T V \end{aligned} \quad (2.1.7)$$

估计误差的均值为：

$$E\{\tilde{x}\} = - (H^T H)^{-1} H^T E\{V\} = 0 \quad (2.1.8)$$

估计误差的均值为零的估计称为无偏估计。因此，如果量测误差是零均值的，则最小二乘估计是无偏估计。

(2) 如果量测误差的均值为零矢量，并设方差为 R ，即：

$$E\{V V^T\} = R$$

则最小二乘估计的均方误差为：

$$E\{\tilde{x} \tilde{x}^T\} = (H^T H)^{-1} H^T R H (H^T H)^{-1} \quad (2.1.9)$$

即估计的均方误差主要决定于量测误差。因为估计无偏，故均方误差也就是误差的方差。

【例 2.1】 用一个仪器对未知的确定标量 x 进行 k 次量测，量测值分别为 $z_1, z_2 \dots z_k$ ，求 x 的最小二乘估计。

【解】 设每次量测的误差为 $V_i (i=1, 2, \dots, k)$ ，则按(2.1.2)式，有：

$$Z = H x + V$$

$$Z^T = [Z_1 \ Z_2 \ \dots \ Z_k], \quad H^T = [1 \ 1 \ \dots \ 1], \quad V^T = [V_1 \ V_2 \ \dots \ V_k]$$

代入(2.1.5)式，得：

$$\hat{x} = \frac{1}{k} (Z_1 + Z_2 + \dots + Z_k)$$

即最小二乘估计就是各次量测的平均值。

$$\text{设: } E\{V\} = 0 \quad E\{V V^T\} = R = \begin{bmatrix} R_1 & 0 \\ & \ddots \\ 0 & R_1 \end{bmatrix}$$

则，估计的均方误差为：

$$E\{\tilde{x}^2\} = \frac{1}{k} R_1$$

这说明量测的次数越多，估计的均方误差越小。

如果 x 为随机变量，要求了解各个时刻的随机值，则每次量测值就是量测时刻的带有量测误差的 x 值。如果仍按上面的计算求最小二乘估计 $\hat{x}$ ，并且认为 $\hat{x}$ 就是各个时刻 x 值的估计，则估计误差既包含量测误差，又包含 x 随机变化的值，误差就要比 x 为确定值的情况

为大, 估计的均方误差亦包含有 x 的方差。因此, 最小二乘估计一般只适用于估计确定值的或有确定性变化的物理量。

【例 2.2】用两个仪器对未知的确定标量 x 各量测一次, 量测值为 Z_1 和 Z_2 , 设仪器的量测误差均值为零, 方差各为 R_1 和 $4R_1$, 求最小二乘估计的均方误差。

$$\text{【解】} \quad H = [1 \quad 1]^T, \quad H^T H = 2, \quad R = \begin{bmatrix} R_1 & 0 \\ 0 & 4R_1 \end{bmatrix}$$

$$E\{\tilde{x}\tilde{x}^T\} = \frac{1}{2} [1 \quad 1] \begin{bmatrix} R_1 & 0 \\ 0 & 4R_1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{2} = \frac{5}{4} R_1$$

即, 用精度相差一倍的两个仪器进行量测, 并按最小二乘估计方法处理量测值, 其效果还没有单独采用一个精度高的仪器好。

2.1.2 加权最小二乘估计

从例 2.2 看出, 最小二乘估计精度不高的原因之一, 是对不同精度的量测值等同对待, 这也是我们不了解量测值统计特性的一种不得已的方法。但是, 如果对不同量测值的精度有一个大致的了解, 即哪一些精确, 哪一些较差, 则可以用加权的方法分别对待, 精度高的加权因子大些, 反之则小些。这就是加权最小二乘估计的思路。它的准则是使:

$$J(\hat{x}) = (Z - H\hat{x})^T W (Z - H\hat{x}) \quad (2.1.10)$$

为最小。式中 W 是适当取值的正定加权矩阵, 当 $W = I$ 时, 上式就是一般最小二乘准则。根据矩阵微分法, 有:

$$\begin{aligned} \frac{\partial J(\hat{x})}{\partial \hat{x}} &= -H^T (W + W^T) (Z - H\hat{x}) \\ &= -2H^T W (Z - H\hat{x}) \end{aligned}$$

令上式为零, 则加权最小二乘估计为:

$$\hat{x} = (H^T W H)^{-1} H^T W Z \quad (2.1.11)$$

估计误差为:

$$\tilde{x} = x - \hat{x} = -(H^T W H)^{-1} H^T W V \quad (2.1.12)$$

如果量测误差 V 的均值为零, 则估计的均方误差为:

$$E\{\tilde{x}\tilde{x}^T\} = (H^T W H)^{-1} H^T W R W H (H^T W H)^{-1} \quad (2.1.13)$$

下面证明: 如果 $W = R^{-1}$, 则估计的均方误差最小。

设 A 和 B 各为 $n \times m$ 和 $m \times l$ 矩阵, 且 AA^T 可逆, 则因:

$$\begin{aligned} & [B - A^T (AA^T)^{-1} AB]^T [B - A^T (AA^T)^{-1} AB] \\ &= B^T B - 2B^T A^T (AA^T)^{-1} AB + B^T A^T (AA^T)^{-1} AA^T (AA^T)^{-1} AB \\ &= B^T B - (AB)^T (AA^T)^{-1} (AB) \\ &\geq 0 \end{aligned}$$

有:

$$B^T B \geq (AB)^T (AA^T)^{-1} (AB) \quad (2.1.14)$$

这就是矩阵型的许瓦茨不等式。

由矩阵知识可知, 正定阵 R 可表示成 $R = S^T S$, 其中 S 为可逆阵。令:

$$\begin{aligned} A &= H^T S^{-1} \\ B &= SWH(H^TWH)^{-1} \end{aligned}$$

有:

$$AB = H^T S^{-1} SWH(H^TWH)^{-1} = I$$

因此:

$$\begin{aligned} (H^TWH)^{-1}H^TWRWH(H^TWH)^{-1} &= B^TB \geq AB)^T(AA^T)^{-1}(AB) \\ &= (AA^T)^{-1} \\ &= (H^TR^{-1}H)^{-1} \end{aligned}$$

即:

$$(H^TWH)^{-1}H^TWRWH(H^TWH)^{-1} \geq (H^TR^{-1}H)^{-1} \quad (2.1.15)$$

当 $W = R^{-1}$ 时, 估计的均方误差最小, 即为 $(H^TR^{-1}H)^{-1}$

$W = R^{-1}$ 的估计称为马尔柯夫估计。即:

$$\hat{x} = (H^TR^{-1}H)^{-1}H^TR^{-1}Z \quad 2.1.16)$$

作这种估计必须预先知道量测误差的方差。

【例 2.3】对例 2.2 采用马尔柯夫估计, 求估计的均方误差。

【解】

$$W = R^{-1} = \begin{bmatrix} R_1 & 0 \\ 0 & 4R_1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{R_1} & 0 \\ 0 & \frac{1}{4R_1} \end{bmatrix}$$

即精度低的量测值加权因子比精度高的小四倍。

$$E\{\hat{x}\hat{x}^T\} = (H^TR^{-1}H)^{-1} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{R_1} & 0 \\ 0 & \frac{1}{4R_1} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}^{-1} = \frac{4}{5}R_1$$

即, 估计的均方误差比精度高的量测值的量测误差方差要小。这说明, 不论仪器的精度多低, 增加不同仪器的量测值总有利于估计精度的提高。

2.1.3 递推最小二乘估计

从以上三个例子看出, 量测值越多, 估计的均方误差就越小。但是这也意味着计算机需要更多的贮存容量来贮存量测值。为了解决这个矛盾, 可采用递推最小二乘估计, 即不需要对所有的量测值一起处理, 而是每次只处理新量测值, 从递推得到新的估计。下面按加权最小二乘估计来推导递推形式的方程。

设根据 k 次量测得到加权最小二乘估计为:

$$\hat{x}_k = (\mathcal{H}_k^T W_k \mathcal{H}_k)^{-1} \mathcal{H}_k^T W_k \mathcal{Z}_k \quad (2.1.17)$$

上式各量的意义都与(2.1.11)式对应的量一样, 只是用下标 k 表示与 k 次量测值有关, 并用草体大写字母表示与第 k 次量测值相区别。例如 $\mathcal{Z}_k$ 表示 k 次量测组成的量测矢量, 以示与第 k 次量测矢量 Z_k 相区别。

令:

$$P_k \triangleq (\mathcal{H}_k^T W_k \mathcal{H}_k)^{-1} \quad (2.1.18)$$

则:

$$\hat{\mathbf{x}}_k = \mathbf{P}_k \mathcal{H}_k^T \mathcal{W}_k \mathcal{Z}_k \quad (2.1.19)$$

同样, 根据 $(k+1)$ 次量测得到的估计为:

$$\hat{\mathbf{x}}_{k+1} = \mathbf{P}_{k+1} \mathcal{H}_{k+1}^T \mathcal{W}_{k+1} \mathcal{Z}_{k+1} \quad (2.1.20)$$

式中:

$$\mathcal{H}_{k+1} = \begin{bmatrix} \mathcal{H}_k \\ \mathbf{H}_{k+1} \end{bmatrix}, \quad \mathcal{W}_{k+1} = \begin{bmatrix} \mathcal{W}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{k+1} \end{bmatrix}, \quad \mathcal{Z}_{k+1} = \begin{bmatrix} \mathcal{Z}_k \\ \mathbf{Z}_{k+1} \end{bmatrix}$$

$$\text{则:} \quad \hat{\mathbf{x}}_{k+1} = \mathbf{P}_{k+1} (\mathcal{H}_{k+1}^T \mathcal{W}_{k+1} \mathcal{Z}_{k+1} + \mathbf{H}_{k+1}^T \mathbf{W}_{k+1} \mathbf{Z}_{k+1}) \quad (2.1.21)$$

$$\begin{aligned} \text{而:} \quad \mathbf{P}_{k+1} &= (\mathcal{H}_{k+1}^T \mathcal{W}_{k+1} \mathcal{H}_{k+1})^{-1} \\ &= (\mathcal{H}_k^T \mathcal{W}_k \mathcal{H}_k + \mathbf{H}_{k+1}^T \mathbf{W}_{k+1} \mathbf{H}_{k+1})^{-1} \\ &= (\mathbf{P}_k^{-1} + \mathbf{H}_{k+1}^T \mathbf{W}_{k+1} \mathbf{H}_{k+1})^{-1} \end{aligned} \quad (2.1.22)$$

根据矩阵求逆公式:

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

可得到:

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \mathbf{P}_k \mathbf{H}_{k+1}^T (\mathbf{H}_{k+1} \mathbf{P}_k \mathbf{H}_{k+1}^T + \mathbf{W}_{k+1}^{-1})^{-1} \mathbf{H}_{k+1} \mathbf{P}_k \quad (2.1.23)$$

由(2.1.19)和(2.1.22)式的关系, (2.1.21)式等号右边第一项可化为:

$$\begin{aligned} \mathbf{P}_{k+1} \mathcal{H}_{k+1}^T \mathcal{W}_{k+1} \mathcal{Z}_{k+1} &= \mathbf{P}_{k+1} \mathbf{P}_k^{-1} \hat{\mathbf{x}}_k \\ &= \mathbf{P}_{k+1} (\mathbf{P}_k^{-1} - \mathbf{H}_{k+1}^T \mathbf{W}_{k+1} \mathbf{H}_{k+1}) \hat{\mathbf{x}}_k \\ &= \hat{\mathbf{x}}_k - \mathbf{P}_{k+1} \mathbf{H}_{k+1}^T \mathbf{W}_{k+1} \mathbf{H}_{k+1} \hat{\mathbf{x}}_k \end{aligned}$$

因此, (2.1.21)式可改写为:

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_k + \mathbf{P}_{k+1} \mathbf{H}_{k+1}^T \mathbf{W}_{k+1} (\mathbf{Z}_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{x}}_k) \quad (2.1.24)$$

(2.1.23)和(2.1.24)式就组成了递推最小二乘估计的一组计算公式。从式中可以看出估计的过程, 即得到新的量测值 $\mathbf{Z}_{k+1}$ 后, 原来的估计 $\hat{\mathbf{x}}_k$ 就要作修正。修正值是新量测值 $\mathbf{Z}_{k+1}$ 与用估计 $\hat{\mathbf{x}}_k$ 得到的量测估计之差, 乘上修正系数矩阵 $\mathbf{P}_{k+1} \mathbf{H}_{k+1}^T \mathbf{W}_{k+1}$, 而系数阵中的 $\mathbf{P}_{k+1}$ 是从 $\mathbf{P}_k$ 计算得来的。即得到一次量测值后作一次计算, 每次计算量不大。随着量测次数的增加, 估计更趋准确。计算公式中仅(2.1.23)式需要求解逆矩阵, 它的阶数是这次量测矢量的维数 m_{k+1} , 一般都有 $m_i \leq n$, 而非递推的估计公式(2.1.11)需要计算 n 阶逆阵。相比之下, (2.1.23)式求逆计算量一般总要比(2.1.11)式的求逆计算量少。

【例 2.4】 在例 2.1 求出 $\hat{\mathbf{x}}_k$ 后, 试求有新量测值 $\mathbf{Z}_{k+1}$ 后的递推最小二乘估计计算公式。

【解】 在不了解仪器每次量测的量测误差方差条件下, 无法采用加权方法。因此, 令:

$$\mathbf{W}_i = \mathbf{I} \quad (i = 1, 2, \dots, k, k+1)$$

在例 2.1 中已经求出 $\hat{\mathbf{x}}_k$ 为:

$$\hat{\mathbf{x}}_k = \frac{1}{k} \sum_{i=1}^k \mathbf{Z}_i$$

根据(2.1.18)和(2.1.23)式, 有:

$$\mathbf{P}_k = \left(\begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \right)^{-1} = \frac{1}{k}$$