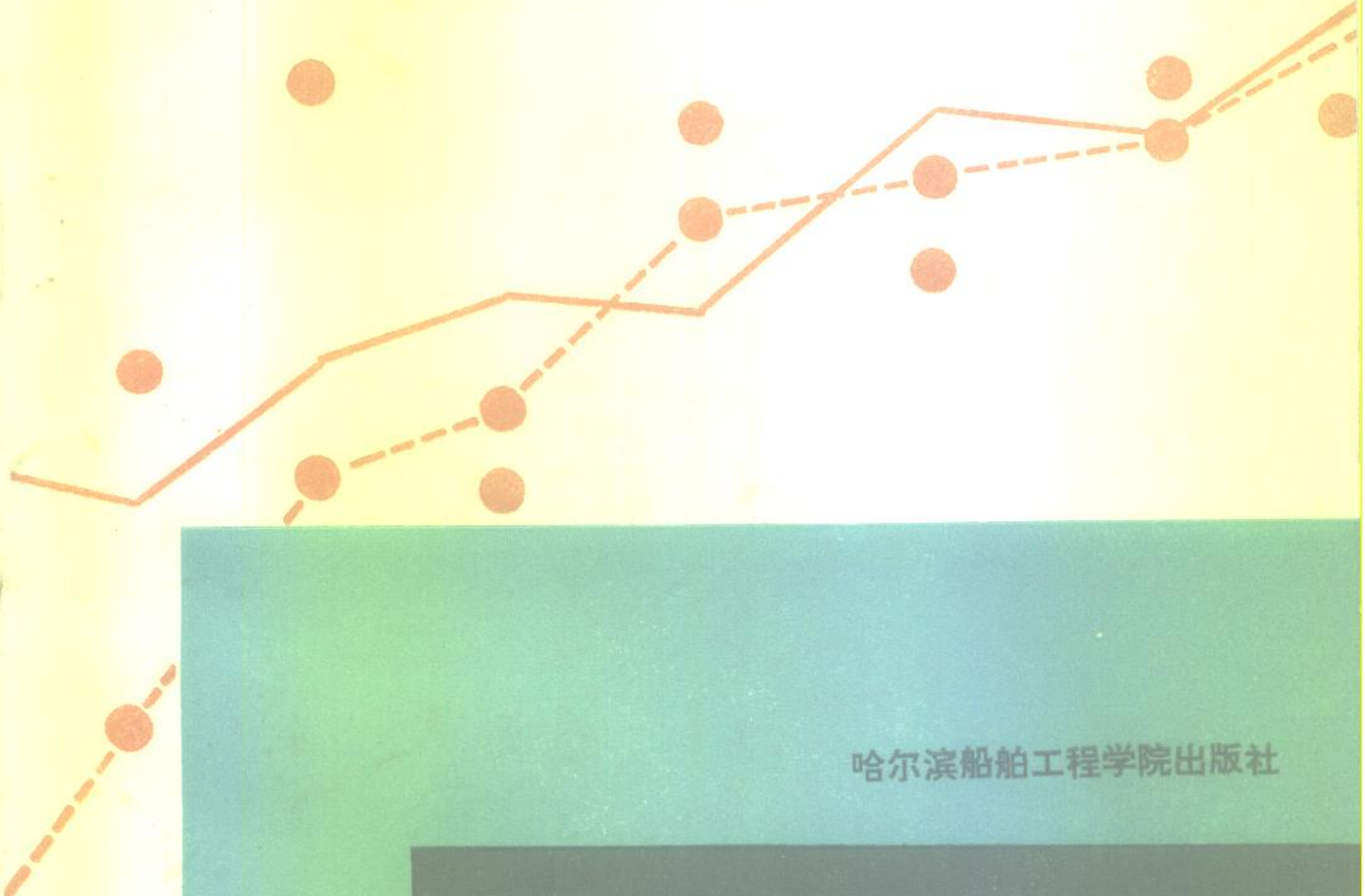

哈尔滨船舶工程学院出版社

离散估计导论

项楚骐　田坦　编著

离散估计导论

项楚骐　田坦　编著

哈尔滨船舶工程学院出版社

内 容 简 介

本书是一本综合介绍估计理论与应用的书，重点放在对离散信号的估计上。全书共分十二章，内容包括确定变量估计与随机变量估计、维纳滤波与自适应维纳滤波、卡尔曼滤波与自适应卡尔曼滤波、线性谱估计与非线性谱估计。

本书可供工科院校的声纳、雷达、通信、控制等专业的研究生、高年级大学生阅读，也可供有关专业科技工作者参考。

离 散 估 计 导 论

项楚骐　田坦　编著

●

哈尔滨船舶工程学院出版社出版

新华书店首都发行所发行

哈尔滨船舶工程学院印刷厂印刷

●

开本　787×1092　1/16　印张21.625　字数508千字

1989年3月第1版　1989年3月第1次印刷

印数：1—1000册

ISBN　7-81007-029-0/O·4

定价：3.60 元

前　言

估计理论与方法的应用十分广泛，不仅在雷达、声纳、控制等传统学科方面有广泛应用，而且在图象处理、功率谱估计、模型参数辨识等新兴学科方面也必不可少。此外，它还渗透到经济计量学、人口控制、市场预测等社会科学领域。由此可见，估计理论与方法在现代科学技术领域有着十分重要的地位与作用。

估计方法的起源可以追溯到二百多年前高斯提出的最小二乘估计方法，后来随着概率论的发展，到本世纪三十年代，出现了贝叶斯估计方法，它们都属于变量估计。对于随机过程的估计则自本世纪四十年代初，维纳提出了适用于平稳随机过程的估计理论——维纳滤波。到六十年代初，卡尔曼把状态变量法应用于线性动态过程的描述，提出了适用于非平稳过程，多输入多输出的递推估计方法——卡尔曼滤波。到六十年代末，人们对随机过程在频率域上的谱分布发生了兴趣，先是提出了线性谱估计方法，随之提出了非线性谱估计方法。近十余年来，各种自适应算法和快速算法的提出，使估计理论及其应用范围达到了一个新的高度。

由于近数十年来提出的估计理论与方法众多，散见于大量文献，本书对此作一归纳。本书的编写遵照以下几个原则：

1. 本书内容按照估计理论的体系进行编排，第一部分为变量估计，第二部分为时间过程估计，第三部分为谱估计，但是注意到这三者之间不是互相孤立的，而是互相渗透和紧密联系的，因而本书采取统一的观点和方法加以阐述。

2. 由于科学技术的发展，大量的快速的观测结果需要处理，并需要借助计算机实施，所以现代估计问题大都以离散方式出现，因而本书着重讨论离散估计。

3. 本书按照工程学观点编写，对于理论不强调数学上的严密性，而注重实用性与简明性，本书中还给出了较多的工程上应用的例子，对于掌握理论部分将是有益的。

全书共分十二章，内容包括确定变量估计与随机变量估计、维纳滤波与自适应维纳滤波、卡尔曼滤波与自适应卡尔曼滤波、线性谱估计与非线性谱估计等。

学习本书需要的基础包括矩阵理论，数字信号处理，概率论与随机过程论的基本内容。本书可供高等工科院校的声纳、雷达、通信、控制等专业的研究生、高年级大学生阅读，也可供有关专业的广大科技工作者参考。

本书的第十一、十二章由田坦执笔，其余由项楚骐执笔。由于我们水平有限，理论与工程素养均感不足，不当之处一定不少，如蒙指正，竭诚欢迎。

编著者

1986年12月

序

在现实世界中，从混杂的环境中提取所需的正确信息来，这是现代信息处理技术的一项主要任务，估计理论是可以在多维领域内实现这种信息处理的一种有力工具，在数学计算和广泛应用的现代，离散估计理论及其应用显得特别受人重视。

作者在本书的撰写中把变量估计、过程估计及谱估计采取统一的观点和方法来加以讨论是一个特点。

本书对于从事声纳、雷达、定位、通信等专业方面的工程技术人员、高年级学生及研究生是一本有用的参考书。

国防科技大学教授、博士生导师
孙 仲 康
1988年1月7日

目　　录

第一篇　变量估计

第二篇　维纳滤波

第三篇 卡尔曼滤波

第四篇 谱 估 计

第一章 绪　论

§1-1 估计与估计对象

什么是估计？简要地说，就是利用测量数据对未知变量或未知时间过程进行推算，推算的方法称为估计方法，推算的结果称为估计。若被估计对象是变量就称为变量估计，若被估计对象是时间过程就称为过程估计。若对时间过程在频率域上的谱分布进行估计则称为谱估计。

当用任何一种方法测量某些物理量（如长度、速度、重量、电压等）时，不可避免地具有随机误差，如进行多次测量，各次读数可能不相同，我们的任务就是要从这些测量值中对物理量作出尽可能好的估计，或者说希望估计值尽可能接近真值，这些物理量的真值是期望得到的，称之谓期望信号。

为了得到好的估计，人们发展了许多种估计方法，然而不论采用何种方法，也不可能得到真值，仍然具有一定的误差。由于采用的估计方法不同，估计的效果也不尽相同。因此需要建立统一标准对估计效果加以衡量，以便评出孰优孰劣。

此外，估计需要按照一定的准则进行，由于所采用准则的不同，将得出不同的估计方法。一种估计方法也只是在某种条件下适用，所以我们将探讨各种估计方法与所应用的准则及适用条件之间的关系。

下面对以上两个问题进行详细说明。

§1-2 估计质量评价

标量情况

我们用符号$\hat{x}$表示对未知物理量的一个估计，$\hat{x}$本身是一个随机变量，它具有一定形状的概率密度函数$p(\hat{x})$，我们希望$\hat{x}$集中在真值x的附近。例如图1-2-1中估计量1要优于估计量2，通常用下面三个指标来评价估计的质量。

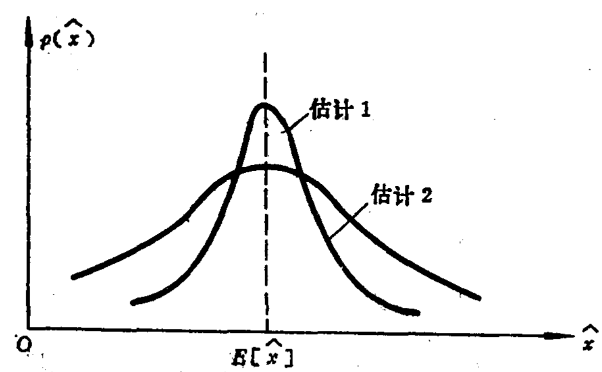

图1-2-1　两个标量估计的比较

1. **偏量**

估计量的均值$E[\hat{x}]$与期望值$\bar{x}$之差定义为此估计值的偏量b，即有

$$b = E[\hat{x}] - \bar{x} = E[\hat{x} - \bar{x}] \tag{1-2-1}$$

当偏量b等于零时，称此估计$\hat{x}$为无偏估计。

若测量数据有限时的估计是有偏的，而当测量数据无限时，偏量趋于零，则称此估计为渐近无偏估计。

2. 方差

估计$\hat{x}$的方差定义为

$$\mathrm{var}[\hat{x}] = E[(\hat{x} - m)^2] = \sigma_{\hat{x}}^2 \tag{1-2-2}$$

$$m = E[\hat{x}]$$

假定$\hat{x}$的概率密度为高斯分布，即有

$$p(\hat{x}) = \frac{1}{\sigma_{\hat{x}}\sqrt{2\pi}} \exp\left[-\frac{(\hat{x} - m)^2}{2\sigma_{\hat{x}}^2}\right] \tag{1-2-3}$$

方差$\sigma_{\hat{x}}^2$决定了密度函数$p(\hat{x})$的分布宽度，它描述了估计值偏离其均值的散度，$\sigma_{\hat{x}}^2$越小，表明估计效果越好。

3. 均方误差

在许多情况下，具有较小偏量的估计却具有较大的方差，反之亦然。这一事实给对于两个估计的评价带来了困难，因此我们定义估计的均方误差为

$$MSE = E[(\hat{x} - \bar{x})^2] = b^2 + \sigma_{\hat{x}}^2 \tag{1-2-4}$$

在均方误差中同时包括偏量和方差。如果随着测量次数增多，使均方误差越来越小，即偏量与方差都趋于零，则称此估计为**一致性估计**

若估计是无偏的，则均方误差就等于方差，均方误差最小与方差最小等价。

矢量情况

例如要估计平面上一点的坐标位置，这是一个二维矢量的估计问题，此二维矢量为

$$\hat{X} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} \tag{1-2-5}$$

这时估计偏量是一矢量

$$\boldsymbol{b} = \begin{bmatrix} \hat{x}_1 - E[\hat{x}_1] \\ \hat{x}_2 - E[\hat{x}_2] \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{1-2-6}$$

这时需要用估计误差的协方差矩阵来表达估计的分散度，此协方差矩阵定义为

$$\boldsymbol{R} = E[(\hat{\boldsymbol{X}} - \boldsymbol{m})(\hat{\boldsymbol{X}} - \boldsymbol{m})^T] \tag{1-2-7}$$

其中 $\boldsymbol{m} = E[\hat{\boldsymbol{X}}]$

为了表达估计的分散度，需要研究误差$\hat{x}_{1\cdot}$、$\hat{x}_{2\cdot}$的联合分布密度。为了简化代数运算，不失一般性，令 $\boldsymbol{m} = 0$，则$\hat{x}_{1\cdot}$、$\hat{x}_{2\cdot}$的联合概率密度函数为

$$p(\hat{\boldsymbol{X}}_{\cdot}) = \frac{1}{2\pi|\boldsymbol{R}|^{1/2}} \exp\left[-\frac{1}{2}\hat{\boldsymbol{X}}_{\cdot}^T \boldsymbol{R}^{-1} \hat{\boldsymbol{X}}_{\cdot}\right] \tag{1-2-8}$$

其中 $\hat{\boldsymbol{X}}_{\cdot} = \begin{bmatrix} \hat{x}_{1\cdot} \\ \hat{x}_{2\cdot} \end{bmatrix}$

记 $$\hat{X}_\cdot^T R^{-1}\hat{X}_\cdot = c^2,\quad c\text{为常数} \tag{1-2-9}$$

展开(1-2-9) 式，得到一椭圆方程，它是二维概率密度函数的一个剖面，现在来求误差矢量$\hat{X}_\cdot$落在此椭圆内的概率。椭圆的面积为

$$A = |R|^{\frac{1}{2}}\pi c^2 \tag{1-2-10}$$

与 c 和 $c+\delta c$ 相对应的两个椭圆之间的微分面积为

$$\delta A = |R|^{\frac{1}{2}} 2\pi c\,\delta c \tag{1-2-11}$$

这个微分面积上概率密度的高度为

$$(2\pi|R|^{\frac{1}{2}})^{-1}\exp\left(-\frac{c^2}{2}\right) \tag{1-2-12}$$

用(1-2-11)乘(1-2-12)式，并进行积分。记落入椭圆内的概率为 P，则在椭圆外的概率为 $1-P$。

$$1-P = \int_c^\infty u\exp\left(-\frac{u^2}{2}\right)du = \exp\left(-\frac{c^2}{2}\right) \tag{1-2-13}$$

$$P = 1-\exp\left(-\frac{c^2}{2}\right) \tag{1-2-14}$$

椭圆方程$\hat{X}_\cdot R^{-1}\hat{X}_\cdot = c^2$称为不确定椭圆或密集椭圆，如图1-2-2所示。

对于任意 k 维矢量，类似结果也成立，我们将在 $k+1$ 维空间中得到一个多维空间的超椭球。

对于图1-2-2 中的不确定椭圆，当$\hat{x}_{1\cdot}$和$\hat{x}_{2\cdot}$不相关时，R^{-1}是对角线阵，不确定椭圆是一正椭圆，而当$\hat{x}_{1\cdot}$和$\hat{x}_{2\cdot}$相关时，R^{-1}是一方阵，不确定椭圆是一斜椭圆。

在矢量情况，均方误差变成了均方程差矩阵$\boldsymbol{P}$。

$$\boldsymbol{P} = E[(\hat{\boldsymbol{X}}-\boldsymbol{X})\times(\hat{\boldsymbol{X}}-\boldsymbol{X})^T]$$

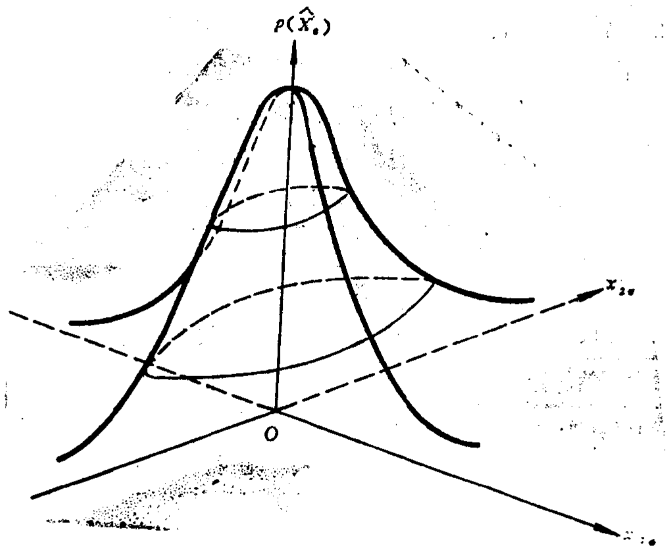

图1-2-2 二维高斯密度函数图形

通常用$\boldsymbol{P}$ 的迹来表达矢量估计的均方误差。

$$\mathrm{tr}\boldsymbol{P} = \mathrm{tr}E[(\hat{\boldsymbol{X}}-\boldsymbol{X})(\hat{\boldsymbol{X}}-\boldsymbol{X})^T]$$

$$= E[(\hat{\boldsymbol{X}}-\boldsymbol{X})^T(\boldsymbol{X}-\boldsymbol{X})]$$

当测量数据无限增多时，$\mathrm{tr}\boldsymbol{P}$ 趋于零，则称此矢量估计为一致性估计。倘若矢量估计是无偏的，则均方误差矩阵就等于协方差矩阵。

§1-3 估计方法名称、符号，应用准则与条件

本书中把各种估计按下表加以分类：

- 估计
  - 变量估计
    - 确定变量估计
    - 随机变量估计
  - 过程估计
    - 维纳滤波
    - 卡尔曼滤波
  - 谱 估 计
    - 线性功率谱估计
    - 非线性功率谱估计

所谓变量估计是这样一类估计问题，根据测量值集合对期望值（标量或矢量）进行估算。当期望值是确定变量时，就称为确定变量估计，当期望值是随机变量时，就称为随机变量估计。

所谓过程估计是根据测量时间序列对期望时间过程进行估算。过程估计方法分成维纳滤波与卡尔曼滤波两大类。维纳滤波是建立在批量数据处理基础之上的，以横向滤波器作为其基本结构形式；卡尔曼滤波是建立在递推数据处理基础之上的，以递归滤波器作为其基本结构形式。

所谓功率谱估计是对随机时间信号在频率域上的功率谱分布进行推算。功率谱估计方法也分成两大类。一类是线性功率谱估计，它是建立在线性付立叶变换的基础之上的；另一类是非线性功率谱估计，它是建立在信号模型辨识的基础之上的。由于谱分布与时间序列之间成非线性变换关系，故称为非线性功率谱估计。

下面对每一种估计方法的名称、符号，应用的准则与使用条件以表格形式列出，本节所述实际上是全书的纲要，也是对各种估计方法的一个归纳。

表1-1 确 定 变 量 估 计

估 计 名 称	符 号	准 则	先验统计知识
最小二乘估计	$\hat{X}_{ls}$	平方误差最小	不需要
最优加权最小二乘估计	$\hat{X}_{lsw}$	加权平方误差最小	测量噪声协方差阵
最大似然估计	$\hat{X}_{ml}$	条件密度$P(Z/X)$最大	条件密度$P(Z/X)$

其中$Z=[z_1,z_2,\cdots,z_N]^T$表示测量集合

表1-2 随 机 变 量 估 计

估 计 名 称		符号	准 则	先验统计知识
贝叶斯估计	最大后验估计	$\hat{X}_{map}$	后验密度$P(X/Z)$最大	$P(X/Z)$已知
	最小方差估计	$\hat{X}_{mv}$	后验密度$P(X/Z)$的均值	$P(X/Z)$已知
线性最小方差估计		$\hat{X}_{lmv}$	均方误差最小 $\hat{X}_{lmv}=A+BZ$	Z和X的头两阶矩
最小均方误差估计		$\hat{X}_{lmse}$	均方误差最小 $\hat{X}_{lmse}=W^TZ$	自相关矩阵ϕ_{ZZ} 互相关矩阵ϕ_{ZX}

表1-3 随机时间过程估计

估计名称	准则	先验统计知识
维纳滤波 （自适应维纳滤波）	均方误差最小	自相关矩阵ϕ_{zz} 互相关矢量ϕ_{zx}
卡尔曼滤波 （自适应卡尔曼滤波）	线性无偏最小方差	状态噪声协方差阵Q_k 测量噪声协方差阵R_k

表1-4 功率谱估计

名称	方法
线性功率谱估计	相关函数法
	周期图法
非线性功率谱估计	时间序列的参数估计法

第一篇 变量估计

第二章 确定变量估计

我们从最基本的估计问题开始，限定被估计量是确定变量。如果对一确定变量进行多次测量，由于测量噪声的存在，各次测量值不相同。因此，需要建立起一套算法，从这些测量值得出确定变量的合理估计。

本章第一节讨论经典的最小二乘估计，第二节讨论最优加权最小二乘估计，后者是前者的改进。第三节讨论非线性最小二乘估计，是线性最小二乘估计在非线性问题中的推广。第四节讨论最大似然估计，最大似然估计通常看成是下章中贝叶斯估计的一个特例，其估计对象仍是确定变量，因而把它列入本章之中。

接着讨论最小二乘数据拟合，所谓数据拟合就是用一确定的数学表达式来近似表示某随机变量与另一确定变量之间的关系，它实际上是最小二乘估计方法的一种应用。最后讨论有约束条件时的最小二乘估计。

§2-1 最小二乘估计$\hat{\mathbf{X}}_{ls}$

最小二乘估计方法是一种最早的估计方法，二百多年前由高斯提出。

标量情况

为了估计未知值x，对它进行了N次测量，测量值是

$$z_i = h_i x + e_i \qquad i = 1, 2, \cdots, N \tag{2-1-1}$$

其中h_i为已知常数，e_i为第i次测量误差。设所求的估计值为$\hat{x}$，则第i次观测值与相应的估计值$h_i\hat{x}$之间的误差为

$$\hat{e}_i = z_i - h_i\hat{x} \tag{2-1-2}$$

将N次误差的平方之和记作

$$J(\hat{x}) = \sum_{i=1}^{N}(z_i - h_i\hat{x})^2 \tag{2-1-3}$$

能使$J(\hat{x})$取极小值的估计值$\hat{x}$就称为未知量x的最小二乘估计，记作$\hat{x}_{ls}$。使$J(\hat{x})$取极小就称为最小二乘准则，根据最小二乘准则求估计值的方法就称为最小二乘估计方法。

采用矩阵符号，可使求最小二乘估计的表达式简练。记

$$\boldsymbol{Z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{pmatrix}, \qquad \boldsymbol{H} = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_N \end{pmatrix}, \qquad \boldsymbol{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{pmatrix} \tag{2-1-4}$$

把(2-1-1)、(2-1-3)式改写成

$$Z = Hx + e \tag{2-1-5}$$

$$J(\hat{x}) = (Z - H\hat{x})^T(Z - H\hat{x}) \tag{2-1-6}$$

求$J(\hat{x})$的极小值，令

$$\frac{\partial J(\hat{x})}{\partial \hat{x}} = -2H^T(Z - H\hat{x}) = 0 \tag{2-1-7}$$

解出最小二乘估计$\hat{x}_{l.s}$

$$\hat{x}_{l.s} = (H^TH)^{-1}H^TZ \tag{2-1-8}$$

由于 $H^TH = h_1^2 + h_2^2 + \cdots + h_N^2 > 0$

显见 $$\frac{\partial^2}{\partial \hat{x}^2}J(\hat{x}) = 2H^TH > 0 \tag{2-1-9}$$

证明了$\hat{x}_{l.s}$是使$J(\hat{x})$达到极小值。

当$h_1 = h_2 = \cdots = h_N = 1$时，容易算出

$$\hat{x}_{l.s} = \frac{1}{N}\sum_{i=1}^{N} z_i \tag{2-1-10}$$

这是N次测量值的算术平均值，所以，算术平均是最小二乘估计的一个特例。

$\hat{x}_{l.s}$既然是一个估计量，它必然是一个随机量，存在着方差。因为估计误差为

$$\begin{aligned} x - \hat{x}_{l.s} &= (H^TH)^{-1}(H^TH)x - (H^TH)^{-1}H^TZ \\ &= (H^TH)^{-1}H^T(Hx - Z) \\ &= -(H^TH)^{-1}H^Te \end{aligned} \tag{2-1-11}$$

假设测量误差矢量e的均值为零，方差阵为R，即

$$E[e] = 0 \qquad E[ee^T] = R$$

则有

$$E[x - \hat{x}_{l.s}] = -(H^TH)^{-1}H^TE[e] = 0$$

可见，当测量误差为零均值时，最小二乘估计为无偏估计。

此时估计误差的方差为

$$\begin{aligned} P_{l.s} &= E[(x - \hat{x}_{l.s})(x - \hat{x}_{l.s})^T] \\ &= E[(H^TH)^{-1}H^T(ee^T)H(H^TH)^{-1}] \\ &= (H^TH)^{-1}H^TRH(H^TH)^{-1} \end{aligned} \tag{2-1-12}$$

若各测量误差独立，且方差相等，即$\operatorname{var} e_i = \sigma^2$，侧得出

$$P_{l.s} = (H^TH)^{-1}\sigma^2 \tag{2-1-13}$$

当$h_1 = h_2 = \cdots = h^N = 1$时

$$P_{l.s} = \frac{1}{N}\sigma^2 \tag{2-1-14}$$

显见，当测量次数N增多时，方差$P_{l.s}$趋近于零。所以，当各测量误差为零均值时，最小二乘估计为无偏一致性估计。

例2-1-1　设用两个电压表测量某电压，第一个表的读数是220伏，已知该表的方差为

8^2（伏^2），第二个表的读数是210伏，已知该表的方差为4^2（伏^2），求估计值$\hat{x}_{L.S}$及其方差。

解：

$$\hat{x}_{L.S}=\frac{220+210}{2}=215\text{伏}$$

$$P_{L.S}=\frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}8^2 & 0\\0 & 4^2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\frac{1}{2}$$

$$=\frac{1}{4}(8^2+4^2)=20\text{伏}^2$$

矢量情况

假设被估计量X是n维矢量，测量Z为m维矢量，则测量矩阵H应是$m\times n$维矩阵，可得出测量方程为

$$\underset{m\times1}{Z}=\underset{m\times n}{H}\ \underset{n\times1}{X}+\underset{m\times1}{e} \tag{2-1-15}$$

仿照标量情况，容易求出最小二乘估计及其方差阵，(2-1-8)式和(2-1-12)式仍然适用，但改成矢量符号

$$\hat{X}_{L.S}=(H^TH)^{-1}H^TZ \tag{2-1-16}$$

$$P_{L.S}=(H^TH)^{-1}H^TRH(H^TH)^{-1} \tag{2-1-17}$$

递推最小二乘估计

如果我们已根据N次测量值求出了最小二乘估计$\hat{X}_{L.N}$及其方差$P_{L.N}$，现在又新增加了一个测量值，必须重新计算一遍才能得出新的估计及其方差，这就需要存储所有测量数据，而且计算量很大，十分不便。因此希望找到一种递推算法，可以由$\hat{X}_{L.N}$和$P_{L.N}$，加上新测量值直接推算出$\hat{X}_{L.(N+1)}$和$P_{L.(N+1)}$。现在来推导这种递推算法。

在有N个测量值时，记测量方程为

$$Z_N=H_NX+\varepsilon_N \tag{2-1-18}$$

X的最小二乘估计为

$$\hat{X}_{L.N}=(H_N^TH_N)^{-1}H_N^TZ_N \tag{2-1-19}$$

现在又得到了一个新测量值

$$Z_{N+1}=h_{N+1}X+e_{N+1} \tag{2-1-20}$$

把(2-1-20)式加到(2-1-18)式中去，得到新的测量方程为

$$Z_{N+1}=H_{N+1}X+\varepsilon_{N+1} \tag{2-1-21}$$

其中

$$Z_{N+1}=\begin{bmatrix}Z_N\\Z_{N+1}\end{bmatrix}\qquad H_{N+1}=\begin{bmatrix}H_N\\h_{N+1}\end{bmatrix}\qquad \varepsilon_{N+1}=\begin{bmatrix}\varepsilon_N\\e_{N+1}\end{bmatrix}$$

由$N+1$个测量值得到的估计应是

$$\hat{X}_{L.(N+1)}=(H_{N+1}^TH_{N+1})^{-1}H_{N+1}^TZ_{N+1} \tag{2-1-22}$$

为了方便起见，记

$$P_N=(H_N^TH_N)^{-1}\quad\text{或}\quad P_N^{-1}=(H_N^TH_N)$$