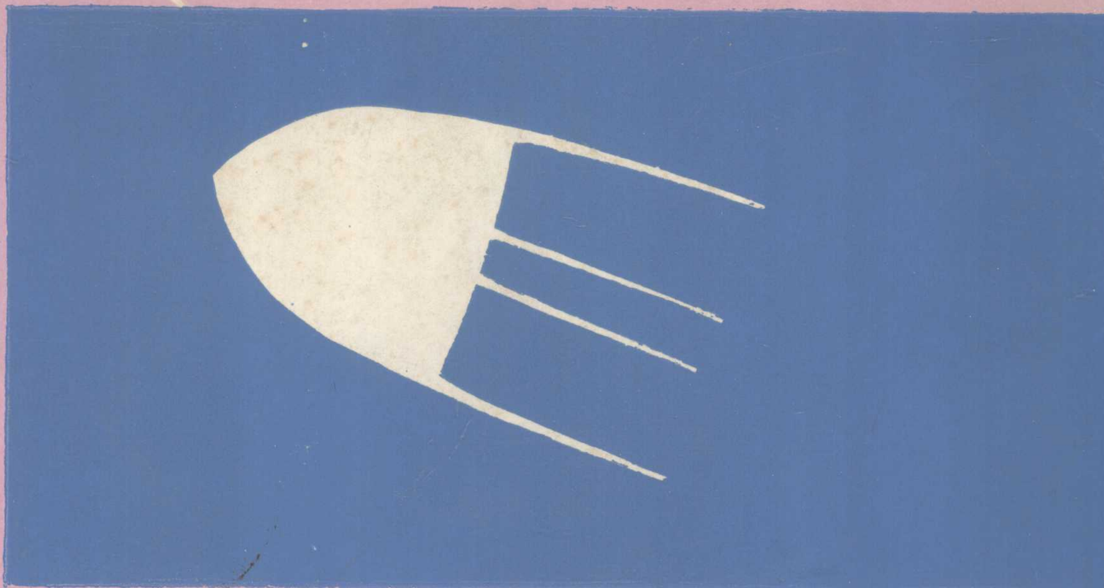
·下 册·



张金槐　常兆诚　编　　吴洪鳌　校

飞行器试验统计学

中国人民解放军国防科学技术大学

飞行器試验統計学

下册

张金槐 常兆诚 编
吴 洪 鳌 校

中国人民解放军国防科学技术大学
1984.5

内 容 提 要

本册（下册）主要讨论线性模型参数估计的 Gauss-Markoff 估计方法、Bayes 方法及线性估计的改进、Kalman 滤波的基本理论及在工程技术中的实现问题、时间序列的统计分析等问题。考虑到工程技术中的需要，列举了在飞行器试验中常被应用的例。

前　　言

这份讲义（下册）是在我们多次举办统计滤波讲座的基础上整理、扩充而成。内容包括线性估计理论、Kalman 滤波的基本理论及滤波技术的实现问题、时间序列的统计分析等。在线性估计问题的理论中，论述了 Gauss-Markoff 的最小二乘方理论、Bayes 估计以及线性估计的改进。我们注意了线性估计的容许性问题。为此，讨论了 Stein 估计及 Minimax 估计。对于超椭球约束下的最小二乘方估计及岭估计，同时进行了讨论。至于 Kalman 滤波，除了给出一般理论外，我们注意了滤波技术的实现问题。因此，比较详细地讨论了自适应滤波方法。至于具体应用，我们只是列举了飞行器试验中常被应用的例，阐明应用的途径。关于时间序列的统计分析，我们只作入门性的介绍，重点是 ARMA 模型参数的估计。至于建模问题只作一般性论述。

在编写这份讲义的过程中，我们得到了国防科工委不少单位的关心和支持。特别是吴洪鳌教授，提出了很多中肯的修改意见，航天工业部 708 所五室的同志还提出了宝贵的意见。在此特致谢意。

在排印过程中，校印刷厂的师付们付出了辛勤劳动，在此一并致谢。

本册的第十章至第十七章由张金槐同志编写，第十八章由常兆诚同志编写。

本讲义适用于飞行器试验统计分析专门化的研究生以及从事试验结果分析的工程技术人员。也可以作为大学本科生有关专业学习统计滤波理论时的参考。

目　录

第十章　线性模型参数估计的最小二乘方理论

前　　言

线性模型未知参数的估计，在飞行器试验统计分析中有着重要的应用。这里所说的估计，主要是线性估计。我们在[1]中讨论过的统计估计方法如最小二乘方(LS)方法，最小方差(MV)估计，最大似然(ML)估计，Bayes 估计以及 Minimax 估计等等均能用于线性模型的参数估计。

对于线性估计，LS 估计已被广泛地应用于工程技术的各个领域。近年来，无偏最小方差估计得到了进一步的运用。然而，从统计决策的观点看，如果用不同的损失函数来衡量估计的优良性，那末古典的最小二乘方方法也还是一种值得商榷的方法。特别是1965年 Jemes-Stein 等人指出，甚至在正态总体的場合，子样的均值并不总是期望值的好的估计。这种观点的提出，使人感到惊讶！因此，估计的容许性 (Admissibility) 研究得到新的重视，而且线性估计的各种改进也就应运而生。我们将会看到，即使是线性估计，也还有不少工作可做。

从这一章开始，我们将对线性模型未知参数的线性估计问题作一个比较系统的论述。大致的內容是线性模型参数的LS估计理论，Bayes 线性估计方法和线性估计的改进问题等。我们先讨论线性模型的最小二乘估计。这方面的工作，不少学者进行了研究([2],[3])。

§1　线性模型及未知参数的最小二乘方估计

设在 t_i 时刻的观测量为 $Z_i, i=1,\cdots,n$，且 Z_i 可表示为未知参数 $\theta_1,\cdots,\theta_m$ 的线性函数和扰动项 ε_i 之和，即

$$Z_i=b_{i1}\theta_1+b_{i2}\theta_2+\cdots+b_{im}\theta_m+\varepsilon_i, i=1,\cdots,n.$$

且假定观测之间是不相关的，$E[\varepsilon_i]=0$，$Var[\varepsilon_i]=\sigma^2, \forall_i$，如果令

$$Z=\begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix},\ X=\begin{bmatrix} \theta_1 \\ \vdots \\ \theta_m \end{bmatrix},\ \mathcal{E}=\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

$$H=[b_{ij}],\ i=1,\cdots,n, j=1,\cdots,m.$$

则

$$Z=HX+\mathcal{E}, \tag{1.1}$$

我们称Z为观测矢量，H为已知的系数矩阵（也称设计矩阵），X为未知的参数矢量，它为常值矢量，$\mathcal{E}$为观测误差矢量。

此处 $E[\mathcal{E}]=0,\quad Var\,\mathcal{E}=E[\mathcal{E}\cdot\mathcal{E}^T]=\sigma^2 I,$

$E[Z]=HX,\quad Var\,Z=\sigma^2 I.$

我们的问题为由观测向量Z去估计未知参数矢量X及σ^2。

(1.1) 为线性估计中最基本的模型，稍为扩充了的模型由 Aitken 考虑过，它为

$$\begin{cases} E[Z]=HX, \\ Var Z=\sigma^2 G, \quad |G|\neq 0. \end{cases} \tag{1.2}$$

模型 (1.2) 表示了观测之间是相关的，其中G假定是已知的，而X，σ^2为未知。易知模型 (1.2) 可以归化为 (1.1)，事实上，作变换

$$Y=G^{-\frac{1}{2}}Z,$$

于是

$$\begin{aligned} E[Y] &= G^{-\frac{1}{2}}HX\triangleq UX, \\ Var Y &= E[(Y-UX)(Y-UX)^T] \\ &= E[G^{-\frac{1}{2}}(Z-HX)(Z-HX)^T(G^{-\frac{1}{2}})^T] \\ &= G^{-\frac{1}{2}}\sigma^2 G(G^{-\frac{1}{2}})^T=\sigma^2 I. \end{aligned}$$

这样就归化为模型 (1.1)。(1.2) 的一个特殊情况为

$$\begin{cases} E[Z]=HX, \\ Var Z=\Sigma, \quad |\Sigma|\neq 0 \end{cases} \tag{1.3}$$

其中Σ为已知。于是在变换$Y=\Sigma^{-\frac{1}{2}}Z$之下，

$$\begin{aligned} E[Y] &= \Sigma^{-\frac{1}{2}}HX=UX, \\ Var Y &= I. \end{aligned}$$

于是 (1.3) 转化为 (1.1)。因此线性模型之下的估计总可以用模型 (1.1) 作为代表来讨论。常称此时之模型为建立在 $(Z, HX, \sigma^2 I)$ 上的线性模型，简记为 $(Z,HX,\sigma^2 I)$。当然上述变换必须在G为非奇异的前提下才成立。如果G为奇异的，则需另作讨论。

线性模型在工程实践中的重要性是不言而喻的。从下面几个例子中可见其一般。

例 1 设

$$E[Z(t)]=\theta_1+\theta_2 t+\cdots+\theta_m t^{m-1},$$

其中 $\theta_1,\cdots,\theta_m$ 是未知待定系数，而在 $t_1,\cdots,t_n$ 所获得的观测值为

$$Z(t_i)\triangleq Z_i=\theta_1+\theta_2 t_i+\cdots+\theta_m t_i^{m-1}+\varepsilon(t_i).$$

引入矢量符号

$$X=[\theta_1\theta_2\cdots\theta_m]^T,$$

$$\mathcal{E}=[\varepsilon(t_1)\cdots\varepsilon(t_n)]^T,$$

$$Z=[Z_1\cdots Z_n]^T,$$

且记

$$H=\begin{bmatrix}1 & t_1 & t_1^2 & \cdots & t_1^{m-1}\\ 1 & t_2 & t_2^2 & \cdots & t_2^{m-1}\\ \vdots & & & & \\ 1 & t_n & t_n^2 & \cdots & t_n^{m-1}\end{bmatrix},$$

则

$$Z=HX+\mathscr{E}.$$

这是熟知的多项式曲线拟合的模型。当然，曲线拟合也可以用其他不同于多项式的函数来进行，比如用

$$E[Z(t)]=\theta_1+\theta_2\sin t+\theta_3 e^{-t}$$

作为拟合曲线，此时X为三維矢量，而H为

$$H=\begin{bmatrix}1 & \sin t_1 & e^{-t_1}\\ 1 & \sin t_2 & e^{-t_2}\\ \vdots & \vdots & \vdots\\ 1 & \sin t_n & e^{-t_n}\end{bmatrix}$$

它为$n\times 3$阶矩阵。

例 2 线性时变离散系统初态$X(0)$的识别问题。

考虑线性离散系统

$$X(k+1)=\Phi(k+1,k)X(k),$$

其中$X(k)$为t_k时刻系统所处的状态。$\Phi(k+1,k)$为转移矩阵。设观测与状态之间由下列关系给出：

$$Z(k)=H(k)X(k)+\varepsilon(k),\ k=1,\cdots,n.$$

于是由

$$X(k)=\Phi(k,0)X(0),$$

且记

$$Z=[Z_1\cdots Z_n]^T,\ \mathscr{E}=[\varepsilon_1\cdots\varepsilon_n]^T,$$

$$\boldsymbol{H}=\begin{bmatrix}H(1)\Phi(1,0)\\ H(2)\Phi(2,0)\\ \vdots\\ H(n)\Phi(n,0)\end{bmatrix},$$

则

$$Z=\boldsymbol{H}\cdot X(0)+\mathscr{E}.$$

由给定之观测矢量Z估计系统之初态$X(0)$，这是熟知的初态识别问题。

例 3 考虑滑动和

$$Y(k)=\sum_{i=1}^{N}W_i\cdot x(k-i),$$

其中$x(k-1),\cdots,x(k-N)$为线性系统的输入序列，$W_i(i=1,\cdots,N)$为权系数，输出信号$Y(k)$的测量值为

$$Z(j)=Y(j)+\varepsilon(j),\ j=k-L,\ k-L+1,\cdots,k,$$

于是由观测$Z(k-L),\cdots,Z(k)$去估计权系数$W_1,\cdots,W_N$的线性模型可写为

$$Z=H\cdot W+\mathscr{E},$$

其中 $Z=[Z(k-L)\quad Z(k-L+1)\cdots Z(k)]^T$,

$$H=\begin{bmatrix} x(k-L-1) & x(k-L-2) & \cdots & x(k-L-N) \\ \vdots & \vdots & \cdots & \\ x(k-2) & x(k-3) & \cdots & x(k-N-1) \\ x(k-1) & x(k-2) & \cdots & x(k-N) \end{bmatrix},$$

$W=[W_1W_2\cdots W_N]^T$,

$\mathcal{E}=[\varepsilon(k-L)\quad \varepsilon(k-L+1)\cdots\varepsilon(k)]^T$

这是 M.Blume 曾讨论过的线性平滑问题([4]). 它仅是线性模型估计中的一个特殊情况而已。

我们再回复至模型 (1.1). 今用最小二乘方法确定出X的估值。为此，作

$$S^2=(Z-HX)^T(Z-HX), \tag{1.4}$$

称 S^2 为残差的平方和，我们取使 $S^2=\min$ 的那种X作为X的估计，且记此估计为 $\hat{X}_{LS}$. 将 (1.4) 式右端关于X求偏导数，且令它为 0，即得

$$\hat{X}_{LS}=(H^TH)^{-1}H^TZ. \tag{1.5}$$

这就是人们熟知的最小二乘方估计，注意到X为m維矢量，Z为n維矢量，H为$n\times m$阶矩阵，因此在 (1.5) 中，$(H^TH)^{-1}$ 为 m阶方阵的求逆。如果H具有最大秩，则 $(H^TH)^{-1}$ 必存在。

由 (1.5) 式可知，如果观测数据增加一个时，则$\hat{X}_{LS}$就要重新计算，这在计算上是不方便的。后面我们将用所谓逐步递推算法来改进这个缺陷。

(1.5) 式为在等精度之下X的最小二乘方估计公式。如果在每次观测中，ε_i的方差不尽相同，例如，记

$$D[\varepsilon_i]=\sigma_i^2,\ i=1,\cdots,n.$$

$$R=\begin{bmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2\end{bmatrix},$$

此时X的最小二乘方估计为使

$$S^2=\|Z-HX\|^2_{R^{-1}}=\min$$

的那种X，我们可知

$$\hat{X}_{LS}=(H^TR^{-1}H)^{-1}H^TR^{-1}Z. \tag{1.6}$$

称上述估计为加权最小二乘方估计，而 R^{-1} 为权系数矩阵。

如果R为正定阵，且H具有最大秩，则 $H^TR^{-1}H$ 的逆矩阵存在。如果 $H^TR^{-1}H$ 为退化的，则最小二乘方解需应用广义逆求解，这种情形将在后面讨论。

最小二乘方估计的性质：

1. 偏倚性(Bias)

如果 $E[\mathcal{E}]=0$，则$\hat{X}_{LS}$为X的无偏估计。

事实上，
$$\begin{aligned}\hat{X}_{LS}&=(H^TR^{-1}H)^{-1}H^TR^{-1}[HX+\mathcal{E}]\\ &=X+(H^TR^{-1}H)^{-1}H^TR^{-1}\mathcal{E},\end{aligned}$$

两边取期望值即得所要的结果。

2. $P \triangleq Var\widetilde{X}=E[(X-\hat{X}_{LS})(X-\hat{X}_{LS})^T]$

$$=(H^TR^{-1}H)^{-1},$$

特殊地，如果 $R=\sigma^2 I$，则

$$P=\sigma^2(H^TH)^{-1}$$

事实上，

$$\begin{aligned}
&E[(X-\hat{X}_{LS})(X-\hat{X}_{LS})^T]\\
&=E\{[(H^TR^{-1}H)^{-1}H^TR^{-1}(HX-Z)]\\
&\quad\cdot[(H^TR^{-1}H)^{-1}H^TR^{-1}(HX-Z)]^T\}\\
&=E\{(H^TR^{-1}H)^{-1}H^TR^{-1}(HX-Z)(HX-Z)^T\\
&\quad[(H^TR^{-1}H)^{-1}H^TR^{-1}]^T\}\\
&=(H^TR^{-1}H)^{-1}H^TR^{-1}E[(HX-Z)(HX-Z)^T]\\
&\quad[(H^TR^{-1}H)^{-1}H^TR^{-1}]^T\\
&=(H^TR^{-1}H)^{-1}H^TR^{-1}\boldsymbol{E}[\mathcal{E}\cdot\mathcal{E}^T]R^{-1}H(H^TR^{-1}H)^{-1},
\end{aligned}$$

由于

$$E[\mathcal{E}\cdot\mathcal{E}^T]=R,$$

于是

$$\begin{aligned}
P&=(H^TR^{-1}H)^{-1}H^TR^{-1}R\cdot R^{-1}H(H^TR^{-1}H)^{-1}\\
&=(H^TR^{-1}H)^{-1}.
\end{aligned}$$

3. 在等精度观测之下，如果H具有最大秩，则 σ^2 的无偏估计为

$$\hat{\sigma}^2=\frac{\widetilde{Z}^T\widetilde{Z}}{\dim[\mathcal{E}]-\dim[X]}=\frac{\widetilde{Z}^T\widetilde{Z}}{n-m}$$

其中 $\widetilde{Z}=Z-\hat{Z}=Z-H\hat{X}_{LS}$, dim [ ] 表示 [ ] 中矢量的維数。

事实上，只需证明

$$E[\widetilde{Z}^T\cdot\widetilde{Z}]=(n-m)\sigma^2$$

就可以了。

注意到

$$\begin{aligned}
Z-H\hat{X}_{LS}&=HX+\mathcal{E}-H[(H^TH)^{-1}H^T(HX+\mathcal{E})]\\
&=HX+\mathcal{E}-H[(H^TH)^{-1}H^THX+(H^TH)^{-1}H^T\mathcal{E}]\\
&=[I_n-H(H^TH)^{-1}H^T]\mathcal{E},
\end{aligned}$$

于是

$$\begin{aligned}
\widetilde{Z}^T\widetilde{Z}&=[(I_n-H(H^TH)^{-1}H^T)\mathcal{E}]^T[(I_n-H(H^TH)^{-1}H^T)\mathcal{E}]\\
&=\mathcal{E}^T(I_n-H(H^TH)^{-1}H^T)^T(I_n-H(H^TH)^{-1}H^T)\mathcal{E},
\end{aligned}$$

注意中间的两个矩阵的乘积为

$$\begin{aligned}
&(I_n-H(H^TH)^{-1}H^T)^T(I_n-H(H^TH)^{-1}H^T)\\
&=(I_n-H(H^TH)^{-1}H^T)^2\\
&=I_n-H(H^TH)^{-1}H^T-H(H^TH)^{-1}H^T+H(H^TH)^{-1}\\
&\quad H^TH(H^TH)^{-1}H^T\\
&=I_n-H(H^TH)^{-1}H^T.
\end{aligned}$$

因此 $I_n - H(H^TH)^{-1}H^T$ 是一个对称，幂等 (idem potent) 矩阵。
于是

$$\widetilde{Z}^T\widetilde{Z} = \mathscr{E}^T[I_n - H(H^TH)^{-1}H^T]\mathscr{E}. \quad (*)$$

再注意到对任意矩阵 $B=(b_{ij})$，有

$$\mathscr{E}^T B \mathscr{E} = \sum_{i=1}^{n} b_{ii}\varepsilon_i^2 + \sum_{i\neq j} b_{ij}\varepsilon_i\varepsilon_j,$$

因此

$$E[\mathscr{E}^T B \mathscr{E}] = \sigma^2\sum_{i=1}^{n} b_{ii} = \sigma^2 \mathrm{tr} B.$$

将上述结果应用于(*)，则有

$$\begin{aligned} E[\widetilde{Z}^T\widetilde{Z}] &= \sigma^2 \mathrm{tr}[I_n - H(H^TH)^{-1}H^T] \\ &= \sigma^2[n - \mathrm{tr}(H(H^TH)^{-1}H^T)]; \end{aligned}$$

$$\begin{aligned} \text{而}\ \mathrm{tr}[H(H^TH)^{-1}H^T] &= \mathrm{tr}[H^T\cdot H(H^TH)^{-1}] \\ &= \mathrm{tr} I_m = m. \end{aligned}$$

这样，

$$E[\widetilde{Z}^T\widetilde{Z}] = \sigma^2(n-m),$$

因此

$$E\left[\frac{\widetilde{Z}^T\widetilde{Z}}{n-m}\right] = \sigma^2.$$

这就证明了所要的结果。

§2 用任意函数的线性组合表示的曲线拟合，正交多项式的运用

这里，对于用任意函数表示的曲线拟合问题作一个简要的讨论。对于 n 个观测值 $Z(t_1),\cdots,Z(t_n)$，我们用下列任意函数的线性组合作为时刻 t 的拟合值

$$Z(t) = \sum_{j=1}^{N}\theta_j\phi_j(t), \quad (2.1)$$

其中 $\theta_1,\cdots,\theta_N$ 为未知参数，$\phi_j(t)$ 为 t 的任意函数（例如，在§1的例1中，$\phi_1=1$，$\phi_2=t$，$\phi_3=t^2,\cdots,\phi_N=t^{N-1}$，它表示了多项式拟合这种特殊情况），记

$$Z = [Z(t_1) \quad Z(t_2) \quad \cdots \quad Z(t_n)]^T,$$

$$H = \begin{bmatrix} \phi_1(t_1) & \phi_2(t_1) & \cdots & \phi_N(t_1) \\ \phi_1(t_2) & \phi_2(t_2) & \cdots & \phi_N(t_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(t_n) & \phi_2(t_n) & \cdots & \phi_N(t_n) \end{bmatrix},$$

$$X = [\theta_1 \cdots \theta_N]^T,$$

则

$$Z = HX + \mathscr{E}.$$

在等精度观测下，X 的 LS 估计为

$$\hat{X}_{LS}=(H^TH)^{-1}H^TZ$$

其中

$$(H^TH)^{-1}=\begin{pmatrix} \sum_{k=1}^{n}\phi_1^2(t_k), & \sum_{k=1}^{n}\phi_1(t_k)\phi_2(t_k), \cdots & \sum_{k=1}^{n}\phi_1(t_k)\phi_N(t_k) \\ \sum_{k=1}^{n}\phi_2(t_k)\phi_1(t_k), & \sum_{k=1}^{n}\phi_2^2(t_k), \cdots & \sum_{k=1}^{n}\phi_2(t_k)\phi_N(t_k) \\ \vdots & & \vdots \\ \sum_{k=1}^{n}\phi_N(t_k)\phi_1(t_k), \cdots\cdots, & & \sum_{k=1}^{n}\phi_N^2(t_k) \end{pmatrix}^{-1}, \quad (2.2)$$

$$H^TZ=\begin{pmatrix} \sum_{k=1}^{n}\phi_1(t_k)Z(t_k) \\ \sum_{k=1}^{n}\phi_2(t_k)Z(t_k) \\ \vdots \\ \sum_{k=1}^{n}\phi_N(t_k)Z(t_k) \end{pmatrix}.$$

由 (2.2) 可以看出，计算逆矩阵是不方便的。但是，容易看出，如果选一个适当的函数基 $\phi_1(t),\cdots,\phi_N(t)$，使它们为离散的标准正交多项式。即是说，它们在 t 的等距的整数点集 1， 2， $\cdots n$ 上为标准正交的，即

$$\begin{cases} \sum_{j=1}^{n}\phi_i(t=j)\phi_l(t=j)=0, i\neq l; \\ \sum_{j=1}^{n}\phi_i^2(t=j)=1. \quad i=l \end{cases}$$

这样，(2.2)中的右端为单位矩阵之逆，而

$$\hat{X}_{LS}=\begin{pmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_N \end{pmatrix}=\begin{pmatrix} \sum_{k=1}^{n}\phi_1(k)Z(k) \\ \vdots \\ \sum_{k=1}^{n}\phi_N(k)Z(k) \end{pmatrix} \quad (2.3)$$

因此运用标准正交多项式后，逆矩阵的计算已不再出现，这在曲线拟合中是最常应用的方法。

关于离散的正交多项式序列，Чебышев， Krautchouk，Charlier，Meixner 以及 Hahn 多项式已由 Abromowitz 和 Stegun[5] 给出，Legendre 和 Laguerre 多项式已由 Morrison[6] 给出。这里，不妨将 Lagendre 正交多项式序列列出于下：

$$\phi_i(\xi)=\frac{p_i(\xi)}{C_i(N)},$$

$$p_i(\xi)=\sum_{\nu=0}^{i}(-1)^\nu C_i^\nu C_{i+\nu}^\nu \frac{\xi(\nu)}{N(\nu)},$$

$$C_i^2(N)=\frac{(N+i+1)\cdot(N+i)\cdots(N+1)}{(2i+1)N(i)},$$

其中

$$C_i^\nu=\frac{i!}{\nu!(i-\nu)!},\quad N(\nu)=N\cdot(N-1)\cdots(N-\nu+1),$$

$$\xi(\nu)=\xi\cdot(\xi-1)\cdots(\xi-\nu+1).$$

$p_i(\xi)$ 的开头几个正交多项式为

i	$p_i(\xi)$
0	1
1	$1-2\frac{\xi}{N}$
2	$1-6\frac{\xi}{N}+6\frac{\xi(\xi-1)}{N(N-1)}$
3	$1-\frac{12\xi}{N}+30\frac{\xi(\xi-1)}{N(N-1)}-20\frac{\xi(\xi-1)(\xi-2)}{N(N-1)(N-2)}$
4	$1-20\frac{\xi}{N}+90\frac{\xi(\xi-1)}{N(N-1)}-140\frac{\xi(\xi-1)(\xi-2)}{N(N-1)(N-2)}$ $+70\frac{\xi(\xi-1)(\xi-2)(\xi-3)}{N(N-1)(N-2)(N-3)}$
⋮	⋮

而对应的 C_i^2 表示为

i	$C_i^2(N)$
0	$N+1$
1	$(N+2)\cdot(N+1)/(3N)$
2	$(N+3)(N+2)(N+1)/[5N(N-1)]$
3	$(N+4)(N+3)(N+2)(N+1)/[7N(N-1)(N-2)]$
4	$(N+5)(N+4)(N+3)(N+2)(N+1)/[9N(N-1)(N-2)(N-3)]$
⋮	⋮

§3 递推最小二乘方法

我们在下列模型之下讨论递推 LS 估计问题：

观测模型为

$$z_i = b_{i1}\theta_1 + b_{i2}\theta_2 + \cdots + b_{im}\theta_m + \varepsilon_i, i = 1, \cdots, n,$$

且

$$E[\varepsilon_i] = 0,$$
$$D[\varepsilon_i] = \sigma_i^2,$$
$$E[\varepsilon_i \varepsilon_j] = 0, i \neq j.$$

记

$$Z_n = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}, \quad X = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_m \end{bmatrix}, \quad \mathscr{E}_n = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

$$H_n = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix},$$

于是

$$Z_n = H_n X + \mathscr{E}_n,$$

且

$$Var\, \mathscr{E}_n \triangleq R_n = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

记 n 次观测后的未知参数 X 的 LS 估计为 $\hat{X}_n$. 假定我们又进行了下一次观测（即第 $n+1$ 次观测），此时我们获得了 z_{n+1}，且

$$z_{n+1} = b_{n+1,1}\theta_1 + b_{n+1,2}\theta_2 + \cdots + b_{n+1,m}\theta_m + \varepsilon_{n+1}, \tag{3.1}$$

记

$$M_{n+1} = [b_{n+1,1} \quad b_{n+1,2} \quad \cdots \quad b_{n+1,m}].$$

则 (3.1) 式可以写为

$$z_{n+1} = M_{n+1} \cdot X + \varepsilon_{n+1}. \tag{3.2}$$

在进行了 $n+1$ 次观测之后，X 的估值 $\hat{X}_{n+1}$ 为

$$\hat{X}_{n+1} = (H_{n+1}^T R_{n+1}^{-1} H_{n+1})^{-1} H_{n+1}^T R_{n+1}^{-1} Z_{n+1}. \tag{3.3}$$

其中

$$H_{n+1} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nm} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n+1,1} & b_{n+1,2} & \cdots & b_{n+1,m} \end{pmatrix} = \begin{pmatrix} H_n \\ \cdots \\ M_{n+1} \end{pmatrix},$$

$$Z_{n+1}=\begin{pmatrix} z_1 \\ \vdots \\ z_n \\ \cdots \\ z_{n+1} \end{pmatrix}=\begin{pmatrix} Z_n \\ \cdots \\ z_{n+1} \end{pmatrix},$$

$$R_{n+1}=\begin{pmatrix} \sigma_1^2 & \cdots\cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots\cdots & \sigma_n^2 & 0 \\ 0 & \cdots & 0 & \sigma_{n+1}^2 \end{pmatrix}=\left(\begin{array}{c|c} R_n & 0 \\ \hline 0 & \sigma_{n+1}^2 \end{array}\right)$$

将上述矩阵和向量的分块表示式代入 (3.3) 式，则

$$\hat{X}_{n+1}=\left(\left[H_n^T \mid M_{n+1}^T\right]\left[\begin{array}{c|c} R_n & 0 \\ \hline 0 & \sigma_{n+1}^2 \end{array}\right]^{-1}\left[\begin{array}{c} H_n \\ \hline M_{n+1} \end{array}\right]\right)^{-1}$$

$$\cdot\left[H_n^T \mid M_{n+1}^T\right]\left[\begin{array}{c|c} R_n & 0 \\ \hline 0 & \sigma_{n+1}^2 \end{array}\right]^{-1}\left[\begin{array}{c} Z_n \\ \hline z_{n+1} \end{array}\right]$$

$$=\left[H_n^T R_n^{-1} H_n+\frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2}\right]^{-1}\left[H_n^T R_n^{-1} Z_n+M_{n+1}^T \sigma_{n+1}^{-2} z_{n+1}\right]$$

或

$$\left(H_n^T R_n^{-1} H_n+\frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2}\right)\hat{X}_{n+1}=H_n^T R_n^{-1} Z_n+\frac{M_{n+1}^T z_{n+1}}{\sigma_{n+1}^2},$$

记
$$\hat{X}_{n+1}=\hat{X}_n+(\hat{X}_{n+1}-\hat{X}_n),$$

于是

$$\left(H_n^T R_n^{-1} H_n+\frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2}\right)\left[\hat{X}_n+(\hat{X}_{n+1}-\hat{X}_n)\right]$$

$$=H_n^T R_n^{-1} Z_n+\frac{M_{n+1}^T z_{n+1}}{\sigma_{n+1}^2},$$

由于

$$(H_n^T R^{-1} H_n)\hat{X}_n=H_n^T R_n^{-1} Z_n,$$

于是

$$\left(H_n^T R_n^{-1} H_n+\frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2}\right)(\hat{X}_{n+1}-\hat{X}_n)$$

$$=-\frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2}\hat{X}_n+\frac{M_{n+1}^T \cdot z_{n+1}}{\sigma_{n+1}^2}$$

$$=\frac{M_{n+1}^T}{\sigma_{n+1}^2}(z_{n+1}-M_{n+1}\hat{X}_n).$$

由上式解出 $\hat{X}_{n+1}$:

$$\hat{X}_{n+1}=\hat{X}_n+\left(H_n^T R_n^{-1} H_n+\frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2}\right)^{-1}\frac{M_{n+1}^T}{\sigma_{n+1}^2}\left(z_{n+1}-M_{n+1}\hat{X}_n\right), \qquad (3.4)$$

上式右端需要对$m \times m$阶矩阵求逆。因此，在未知参数向量X的維数增大时，计算是不方便的。为此，我们运用下述引理（参见附录）

矩阵求逆引理：如果矩阵P_1, P_2, H, R满足方程

$$P_2^{-1}=P_1^{-1}+H^T R^{-1} H, \qquad (3.5)$$

其中，P_1, R为非奇异矩阵，而H具有最大秩，则

$$P_2=P_1-P_1 H^T\left(H P_1 H^T+R\right)^{-1} H P_1 \qquad (3.6)$$

应用矩阵求逆引理的好处在于使逆矩阵计算中的維数降低，这一点在下面的讨论中就会看出来。

矩阵求逆引理有各种其他的形式。例如

$$(A+X^T Y)^{-1}=A^{-1}-A^{-1} X^T\left(1+Y A^{-1} X^T\right)^{-1} Y A^{-1},$$

其中A为方阵，X，Y为行矢量（参见附录）。

有了上述矩阵求逆的引理，我们回复至公式(3.4)

今记

$$P_n=\left(H_n^T R_n^{-1} H_n\right)^{-1},$$

则

$$P_{n+1}=\left(H_{n+1}^T R_{n+1}^{-1} H_{n+1}\right)^{-1}$$

$$=\left(\left[H_n^T \vdots M_{n+1}^T\right]\left[\begin{array}{c|c} R_n & 0 \\ \hline 0^T & \sigma_{n+1}^2 \end{array}\right]^{-1}\left[\begin{array}{c} H_n \\ \hline M_{n+1} \end{array}\right]\right)^{-1}$$

$$=\left(H_n^T R_n^{-1} H_n+\frac{M_{n+1}^T \cdot M_{n+1}}{\sigma_{n+1}^2}\right)^{-1},$$

应用矩阵求逆引理，令

$$P_1=\left(H_n^T R_n^{-1} H_n\right)^{-1}=P_n,$$

$$H=M_{n+1},$$

$$R^{-1}=\frac{1}{\sigma_{n+1}^2}，\text{即}\quad R=\sigma_{n+1}^2,$$

于是

$$P_{n+1}=P_n-P_n M_{n+1}^T\left(\sigma_{n+1}^2+M_{n+1} P_n M_{n+1}^T\right)^{-1} M_{n+1} P_n,$$

而

$$\hat{X}_{n+1}=\hat{X}_n+P_{n+1} \cdot \frac{M_{n+1}^T}{\sigma_{n+1}^2}\left(z_{n+1}-M_{n+1} \hat{X}_n\right).$$

将P_{n+1}的表达式代入上式之$P_{n+1} \cdot M_{n+1}^T \cdot \frac{1}{\sigma_{n+1}^2}$，则它可表示为