自适应控制讲义

第一章 系统辨识

上海师大 袁震東

一九七九年六月

序　言

控制论早期的发展是受到生物学和医学的启示的，在今天，生物学研究仍然给控制论的发展以启示和借鉴。

人是高等哺乳动物，人与其他动物一样具有适应自然环境的能力。举个容易体验的例子来说，人从光线强的地方突然走进一个光线微弱的房间，开始一下子什么也看不见，但过了一会儿眼睛就会适应这种环境看清楚室内的摆设和事物。这是一种自适应现象。

这种自适应现象给控制理论工作者以很大的启示：工业生产过程是变化的，它的环境、条件无不随着时间的推移而变化。如果能够设计一种具有适应能力的控制器来适应这种环境的变化，那么其控制的效果无疑会大大地增强。于是自适应控制就应运而生了。

目前关于自适应方法的研究很多，在这本讲义里我们重点介绍应用广泛、易于普及推广的一种自适应控制器——自校正调节器。

自校正调节器是1970年Peteska提出的，1973年Åström和Witlenmark作了理论研究和应用。目前从理论研究方面，自校正调节器已发展为自校正控制器，自校正预测器，自校正滤波器等并得到了自校正调节器的许多优良性质。从应用方面来看，自校正调节器除了最早用于造纸机的控制外，近几年来在纸浆，矿石粉碎，水泥，化工，超级油轮驾驶，二氧化钛窑的控制都取得了成功。在现代控制论中得到如此广泛应用的系统确实还是不多见的。

自校正调节器本质上是一个预测器，在某些条件下它收敛于最小方差控制器，在自校正调节器中参数是在线辨识的，因此在本讲义的第一章介绍了最小二乘法辨识方法，这是一种简单易行的方法。

由于这本讲义是笔者在参加各种学术活动的间歇中写成的，可能不够连贯，同时还有错误，希望读者给予批评指正。

作者

1979.6.16.

第一章　系统辨识

§1　引　言

从输入、输出的量测出发建立对象及其干扰的数学模型是计算机控制的必要前提。

从物理或化学的机理来建立对象及其干扰的数学模型固然是一种办法。但实际使用起来，由于建立模型时略去了许多因素，往往使得模型不能真实反映系统动态。而且这样建立的模型往往过于复杂，难于进行系统的分析和设计。从输入，输出数据建立具有一定精度的数学模型反有着目的性强、简单易行等许多优点。因此近十几年来，时间序列，系统辨识受到愈来愈多的重视。

系统辨识的主要内容包括：系统辨识的试验设计，模型结构的确定，参数估计方法，模型验证等诸方面。本章不能面面俱到介绍，只系统阐述瑞典控制理论专家 Åström 关于最小二乘法辨识的理论和方法。这是一个完全、简单、易行的方法。

§2　什么叫系统辨识

如上所述，系统辨识理论是利用试验或运行数据构造数学模型的理论。在许多场合下都希望建立具有一定精度的数学模型。例如，设计计算机控制系统要建立生产过程和对象的数学模型；研究生物控制系统要建立生物系统的模型；研究经济学要研究社会经济的模型等等。

关于系统辨识 Zadeh[1] 曾经给出一个定义："系统辨识是在输入、输出的基础上，从一类系统中确定一个与所测系统等价的系统"。

这里的等价并非代数等价。它的意义可以说明如下：在系统辨识时，我们要建立一个准则用以表示系统的模型与实测值的不符情况。一般地说这个准则（也称损失函数）是过程输出和模型输出的函数：

$$V = V(y.y_m)$$

其中 y 是过程输出（即实际输出），y_m 是模型输出．

在系统辨识中，如果对于两个模型 m_1 和 m_2 具有相等的损失函

数，即

$$V(y, y_{m_1}) = V(y, y_{m_2})$$

那么，我们称m_1和m_2是等价的。用损失函数来定义等价时，辨识问题成了一个最优化问题：找出一个模型

$$S_0 \in \mathcal{G} \qquad \mathcal{G}\text{是我们考察的模型类}$$

使得

$$V(S_0) = \min_{S \in \mathcal{G}} V(y, y_m)$$

如果$V(S)$存在唯一个使得它达到最小值。那么称模型类$\mathcal{G}$被辨识了。

在这一章里，我们主要讨论线性系统的辨识，如所周知一个定常线性系统其表示方式可以各式各样，对于各种不同的表示方式，有着不同的辨识方法。

从系统输入、输出，确定定常线性系统的脉冲响应，可用下一节讨论的相关分析方法。

为了确定线性系统差分方程的系数和阶可以用§3中讨论的最小二乘法。

例1，考察一个单输入一个单输出，离散线性系统，假定把一个输入序列$\{u(1), u(2), \cdots, u(N)\}$，作用于下列定常线性系统。

$$y(t) + ay(t-1) = bu(t-1) \qquad (1-1)$$

测得系统的输出$\{y(1), y(2), \cdots, y(N)\}$。辨识的目的是确定模型（1—1）中的$a$、$b$，

在这里(1-1)式表示了一个模型类，确定a、b就是确定模型类中哪一个模型与我们实际界测的模型等价。损失函数可取为

$$V(y, y_m) = \sum_{t=2}^{N} e^2(t)$$

这里

$$e(t) = y(t) + ay(t-1) - bu(t-1) \qquad (1-2)$$

例2．考察一个时间离散的干扰．假定这个干扰的观测值为序列$\{y(1), y(2), \cdots, y(N)\}$，干扰序列的模型为

$$y(t) + ay(t-1) = 0$$

损失函数取作

$$V(y, y_m) = \sum_{t=2}^{N} e^2(t)$$

其中

$$e(t) = y(t) + a y(t-1)$$

这类问题都可以归结为§3中所讨论的最小二乘法问题。

对于同一个输入—输出或者相同的外部特征，它可以对应无数个模型。两个定常线性系统，如果它们的传递函数相同，那么我们称它们是K—等价的。K—等价是一种数学的等价关系，它满足等价关系的三个条件：

(i) 对$\mathcal{S}$的任一元素S，$S \overset{K}{\sim} S$（S和它自身是K等价的）；

(ii) 对于S_1，$S_2 \in \mathcal{S}$，如果$S_1 \overset{K}{\sim} S_2$ 那么$S_2 \overset{K}{\sim} S_1$；

(iii) 如果S_1，S_2，$S_3 \in \mathcal{S}$且$S_1 \overset{K}{\sim} S_2$，$S_2 \overset{K}{\sim} S_3$ 那么必定有$S_1 \overset{K}{\sim} S_3$ 成立。

在$\mathcal{S}$上定义了等价关系，就可以在$\mathcal{S}$上定义等价类。所谓等价类$\varphi \subset \mathcal{S}$，有$S_1 \overset{K}{\sim} S_2$，$S_1 \in \varphi$ 可推出$S_2 \in \varphi$，反过来若$S_1 \overset{K}{\sim} S_2$且$S_2 \in \varphi$也可推出$S_1 \in \varphi$。

全体等价类$\{\varphi\}$所成的集合称为商集，记作$\mathcal{S}/\sim$．从系统输入输出量测，我们只能希望指出商集中的哪一个元素与数据相符而不能希望区分等价类中的不同元素。因此系统辨识涉及到系统结构，规范形，参量化等许多问题。这里无法作更详细的介绍，有兴趣的读者可参见。[2][3][4]

§3　系统辨识的最小二乘法[5]

十九世纪初由于航海事业发展的需要，在研究天体运动轨迹计算时高斯和勒让特各自独立地开创了最小二乘法。二十世纪六十年代后期，Åström 等人把这个方法用于动态系统辨识，获得了许多成果。目前，最小二乘法已成为系统辨识的一种基本的方法。

§3.3.1　最小二乘法估计公式

设考虑一个单输入单输出的定常线性系统。如果采样间隔相等，

输入信号在每一个采样间隔等于常数，即输入是一个按段常数函数，那么这个系统一般可表示为

$$y(t)+a_1y(t-1)+\cdots+a_ny(t-n)=b_1u(t-1)+\cdots+b_nu(t-n)+e(t) \qquad (3-68)$$

如果我们用一个变化的输入信号 $u(t)$，记录各时间间隔的输入和相应的输出值，得到一个输入输出对序列 $\{(u(t),y(t)),\ t=1,2,\cdots,n+N\}$

（3－68）模型包括了（1－1），（1－2）给出的模型，也概括了差分方程模型，辨识（1－68）给出的模型类包括二个问题：

1. 在已知模型阶 n 时，如何从输入输出估计模型中的未知参数 $a_1,\cdots,a_n,b_1,\cdots,b_n$；

2. 如何从输入输出确定模型的阶 n。

下面我们用最小二乘法来介第 1 个问题。第 2 个问题则放到§3.3.5节中加以介决。

在最小二乘法中，损失函数取作

$$V(y,y_m)=\sum_{t=n+1}^{N+n} e^2(t) \qquad (3-69)$$

即

$$V(y,y_m)=\sum_{t=n+1}^{n+N}[y(t)+a_1y(t-1)+\cdots+a_ny(t-n)-b_1u(t-1)\cdots-b_nu(t-n)]^2$$

对于给出的 $(u(t),y(t))$ $t=1,2,\cdots,n+N$，如果能找到一组参数 $\hat{a}_1,\hat{a}_2,\cdots,\hat{a}_n,\hat{b}_1,\hat{b}_2,\cdots,\hat{b}_n$ 使得（3－69）达到最小值，则称 $\hat{a}_1,\hat{a}_2,\cdots,\hat{a}_n,\hat{b}_1,\hat{b}_2,\cdots,\hat{b}_n$ 是参数 $a_1,a_2,\cdots,a_n,b_1,b_2,\cdots,b_n$ 的最小二乘估计。

为了清楚地叙述最小二乘估计的求介过程，我们引进如下的向量和矩阵

$$Y=\begin{pmatrix} y_{n+1}\\ y_{n+2}\\ \vdots\\ y_{n+N}\end{pmatrix},\quad e=\begin{pmatrix} e_{n+1}\\ e_{n+2}\\ \vdots\\ e_{n+N}\end{pmatrix},\quad \theta=\begin{pmatrix} a_1\\ a_2\\ \vdots\\ a_n\\ b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix} \qquad (3-70)$$

$$
\Phi = \begin{pmatrix} -y_n & -y_{n-1} & \cdots & -y_1 & u_n & u_{n-1} & \cdots & u_1 \\ -y_{n+1} & -y_n & \cdots & -y_2 & u_{n+1} & u_n & \cdots & u_2 \\ \vdots & & & & & & & \vdots \\ -y_{n+N-1} & -y_{n+N-2} & \cdots & -y_N & u_{n+N-1} & u_{n+N-2} & \cdots & u_N \end{pmatrix} \quad (3-71)
$$

其中 y_i 表示 $y(i)$ u_i 表示 $u(i)$，e_i 表示 $e(i)$，$i=1,2,\cdots,n+N$.

利用向量和矩阵记号（3－68）可以写成

$$
y = \Phi\theta + e \quad (3-72)
$$

损失函数为

$$
V = \sum_{t=n+1}^{N} e_i^2 = e^T e \quad (3-73)
$$

最小二乘法是找 θ 的估计值 $\hat{\theta}$ 使得（3－73）达到最小。

因为

$$
\begin{aligned}
e^T e &= (y - \Phi\theta)^T (y - \Phi\theta) \\
&= y^T y - y^T \Phi\theta - \theta^T \Phi^T y + \theta^T \Phi^T \Phi\theta
\end{aligned}
$$

假定 $\Phi^T\Phi$ 非奇异，那么 $\Phi^T\Phi$ 必为正定矩阵，我们把上式配成完全平方的形式得

$$
\begin{aligned}
e^T e &= \{\theta - (\Phi^T\Phi)^{-1}\Phi^T y\}^T \Phi^T\Phi \{\theta - (\Phi^T\Phi)^{-1}\Phi^T y\} \\
&\quad + y^T y - y^T \Phi (\Phi^T\Phi)^{-1} \Phi^T y \\
&\geqslant y^T y - y^T \Phi (\Phi^T\Phi)^{-1} \Phi^T y
\end{aligned}
$$

当

$$
\theta - (\Phi^T\Phi)^{-1}\Phi^T y = 0
$$

时上述不等式中等号成立。

因此 θ 的最小二乘估计为

$$
\hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^T y \quad (3-74)
$$

（3－74）是最小二乘估计的一般公式，Φ 和 y 都是由 $(u(t), y(t))$，$t=1,2,\cdots,n+N$ 所给出的，应用（3－74）就可以求得 θ 的最小二乘估计。

例1，考虑§3.1节中例1给出的模型。

$$y(t)+ay(t-1)-bu(t-1)=e(t)\qquad t=1,2,\cdots,N+1$$

由(3-70)，(3-71)

$$Y=\begin{pmatrix} y_2\\ y_3\\ \vdots\\ y_{N+1}\end{pmatrix},\quad \Phi=\begin{pmatrix} -y_1 & u_1\\ -y_2 & u_2\\ \vdots & \vdots\\ -y_N & u_N\end{pmatrix},\quad e=\begin{pmatrix} e_2\\ e_3\\ \vdots\\ e_{N+1}\end{pmatrix},\quad \theta=\begin{bmatrix} a\\ b\end{bmatrix}$$

上式可以写成

$$Y=\Phi\theta+e$$

计a

$$\Phi^T\Phi=\begin{bmatrix} \sum_{i=1}^{N} y_i^2 & -\sum_{i=1}^{N} y_i u_i\\ -\sum_{i=1}^{N} u_i y_i & \sum_{i=1}^{N} u_i^2\end{bmatrix}$$

求$\Phi^T\Phi$的逆矩阵得

$$(\Phi^T\Phi)^{-1}=\frac{1}{\left(\sum_{i=1}^{N} u_i^2\right)\left(\sum_{i=1}^{N} y_i^2\right)-\left(\sum_{i=1}^{N} y_i u_i\right)^2}\begin{bmatrix} \sum_{t=1}^{N} u_i^2 & +\sum_{i=1}^{N} u_i y_i\\ +\sum_{i=1}^{N} u_i y_i & \sum_{i=1}^{N} y_i^2\end{bmatrix}$$

又

$$\Phi^T Y=\begin{bmatrix} -\sum_{i=2}^{N+1} y_i y_{i-1}\\ \sum_{i=2}^{N+1} y_i u_{i-1}\end{bmatrix}=\begin{bmatrix} -\sum_{i=1}^{N} y_{i+1} y_i\\ \sum_{i=1}^{N} y_{i+1} u_i\end{bmatrix}$$

由(3-74)公式得

$$\hat{a}=\frac{-\left(\sum_{i=1}^{N} u_i^2\right)\sum_{i=1}^{N} y_i y_{i+1}+\left(\sum_{i=1}^{N} u_i y_i\right)\sum_{i=1}^{N} u_i y_{i+1}}{\left(\sum_{i=1}^{N} u_i^2\right)\left(\sum_{i=1}^{N} y_i^2\right)-\left(\sum_{i=1}^{N} u_i y_i\right)^2}$$

$$\hat{b}=\frac{\left(\sum_{i=1}^{N}y_i^2\right)\left(\sum_{i=1}^{N}y_{i+1}u_i\right)-\left(\sum_{i=1}^{N}y_iy_{i+1}\right)\left(\sum_{i=1}^{N}y_iu_i\right)}{\left(\sum_{i=1}^{N}u_i^2\right)\left(\sum_{i=1}^{N}y_i^2\right)-\left(\sum_{i=1}^{N}u_iy_i\right)^2}$$

§3.3.2 关于观察个数的递推

在生产运行过程中，系统的轨入，轨出往往是序贯地测得的，如果我们测得了 $n+N$ 个数据，可以用（3－74）估计出系统所含的未知参数，增加了一个新的数据（即测得了 $n+N+1$ 个数据），那么我们要用（3－74）重新计秝一遍才能得到 $n+N+1$ 个数据时的最小二乘估计.这样，为了记忆 $n+N+1$ 个数据需要占用大具内存，而且许多计秝是重复的。因此使我们想到能否用 $n+N$ 个数据所得的最小二乘估计 $\hat{\theta}(N)$，在获得 $N+n+1$ 个数据时作适当的修改来获得 $n+N+1$ 个数据时的最小二乘估计呢？这就是本节所讨论的递推最小二乘估计。

使用递推最小二乘估计可以减少内存和计秝昇，使得系统的参数随时间变化时用它来及时修正参数。因此最小二乘估计已成为自适应控制的重要工具。

为了导出 θ 的递推估计公式，我们把观测数中的 N 作为变动参数引进下列的符号。

由 $n+N$ 个数据（对）获得的最小二乘估计记作 $\hat{\theta}(N)$，$n+N+1$ 个数据获得的估计记作 $\hat{\theta}(N+1)$。

$$\Phi(N)=\begin{pmatrix}-y_n, -y_{n-1} & \cdots & -y_1 & u_n & u_{n-1} & \cdots & u_1\\ -y_{n+1} & -y_n & \cdots -y_2 & u_{n+1} & u_n & \cdots & u_2\\ \vdots & & & & & & \\ -y_{n+N-1} & -y_{n+N-2} & \cdots -y_N & u_{n+N-1} & u_{n+N-2} & \cdots & u_N\end{pmatrix} \quad (3-75)$$

$$Y(N)=\begin{pmatrix}y_{n+1}\\ y_{n+2}\\ \vdots\\ y_{n+N}\end{pmatrix} \quad (3-76)$$

当增加一个观测对(u_{n+N+1}, y_{n+N+1})时，$\Phi(N)$就增加一行，$Y(N)$增加了一个分量。这时

$$\Phi(N+1)=\begin{bmatrix}\Phi(N)\\ ----\\ \varphi(N+1)\end{bmatrix}, \quad Y(N+1)\begin{bmatrix}Y(N)\\ \\ y_{N+n+1}\end{bmatrix} \tag{3-77}$$

其中

$$\varphi(N+1)=(-y_{n+N}, -y_{n+N-1}, \cdots, -y_{N+1}, u_{N+n}, u_{n+N-1}, \cdots, u_{N+1}) \tag{3-78}$$

观测数为 $n+N$ 时由(3-74)，θ的最小二乘估计为

$$\hat{\theta}(N)=[\Phi^T(N)\Phi(N)]^{-1}\Phi^T(N)Y(N) \tag{3-79}$$

观测数为 $n+N+1$时，θ的最小二乘估计为

$$\begin{aligned}\hat{\theta}(N+1)&=[\Phi^T(N+1)\Phi(N+1)]^{-1}\Phi^T(N+1)Y(N+1)\\ &=[\Phi^T(N)\Phi(N)+\varphi^T(N+1)\varphi(N+1)]^{-1}[\Phi^T(N)Y(N)+\varphi^T(N+1)y_{n+N+1}]\end{aligned} \tag{3-80}$$

为了推导过程中书写方便，在下列的推导过程中暂时省略$\Phi(N)$中的N，记作Φ，省略$\varphi(N+1)$中的$N+1$记作φ，这样做不会引起混淆，推导完以后我们再把省略的N和N+1添上就获得明晰的递推公式，经过省略(3-80)可写成

$$\hat{\theta}(N+1)=(\Phi^T\Phi+\varphi^T\varphi)^{-1}(\Phi^T Y+\varphi^T y_{N+n+1}) \tag{3-81}$$

这样每一次递推计算必须计算$(\Phi^T\Phi+\varphi^T\varphi)$矩阵的逆，这样做工作量较大，为了简化计算我们运用了下列著名的矩阵求逆引理。

<u>求逆引理</u>　　设A、C和$A+BCD$都是非异方阵，

那么

$$(A+BCD)^{-1}=A^{-1}-A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1} \tag{3-82}$$

<u>证明</u>：　用直接代入法证

$$\begin{aligned}&(A+BCD)[A^{-1}-A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}]\\ &=I-B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}+BCDA^{-1}\end{aligned}$$

$$-BCDA^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}$$

$$=I+BCDA^{-1}-B(I+CDA^{-1}B)(C^{-1}+DA^{-1}B)^{-1}DA^{-1}$$

$$=I+BCDA^{-1}-BC((C^{-1}+DA^{-1}B)(C^{-1}+DA^{-1}B)^{-1}DA^{-1}$$

$$=I+BCDA^{-1}-BCDA^{-1}$$

$$=I$$

证毕。

注：当 $D=B^T$ 时有

$$(A+BCB^T)^{-1}=A^{-1}-A^{-1}B(C^{-1}+B^TA^{-1}B)^{-1}B^TA^{-1} \tag{3-83}$$

把（3－83)用于（3－81），此时

$$A=\Phi^T\Phi,\quad B=\varphi^T,\quad C=1$$

而

$$(\Phi^T\Phi+\varphi^T\varphi)^{-1}=(\Phi^T\Phi)^{-1}-(\Phi^T\Phi)^{-1}\varphi^T(1+\varphi(\Phi^T\Phi)^{-1}\varphi^T)^{-1}\cdot\varphi(\Phi^T\Phi)^{-1} \tag{3-84}$$

把（3－84）代入（3－81)得

$$\begin{aligned}\hat{\theta}(N+1)&=(\Phi^T\Phi+\varphi^T\varphi)^{-1}(\Phi^Ty+\varphi^Ty_{n+N+1})\\&=\{(\Phi^T\Phi)^{-1}-(\Phi^T\Phi)^{-1}\varphi^T(1+\varphi(\Phi^T\Phi)^{-1}\varphi(\Phi^T\Phi)^{-1}\}\\&\quad\cdot(\Phi^Ty+\varphi^Ty_{n+N+1})\\&=(\Phi^T\Phi)^{-1}\Phi^Ty+(\Phi^T\Phi)^{-1}\varphi^Ty_{n+N+1}\\&\quad-(\Phi^T\Phi)^{-1}\varphi^T[1+\varphi(\Phi^T\Phi)^{-1}\varphi^T]^{-1}\varphi(\Phi^T\Phi)^{-1}\Phi^Ty\\&\quad-(\Phi^T\Phi)^{-1}\varphi^T[1+\varphi(\Phi^T\Phi)^{-1}\varphi^T]^{-1}\varphi(\Phi^T\Phi)^{-1}\varphi^Ty_{n+N+1}\\&=\hat{\theta}(N)+(\Phi^T\Phi)^{-1}\varphi^T[1+\varphi(\Phi^T\Phi)^{-1}\varphi^T]^{-1}[y_{n+N+1}-\varphi\hat{\theta}(N)]\end{aligned}$$

在上式中令

$$K(N)=(\Phi^T\Phi)^{-1}\varphi^T[1+\varphi(\Phi^T\Phi)^{-1}\varphi^T]^{-1} \tag{3-85}$$

得

$$\hat{\theta}(N+1)=\hat{\theta}(N)+K(N)[y_{n+N+1}-\varphi\hat{\theta}(N)] \tag{3-86}$$

在(3-85)中，$(\Phi^T\Phi)^{-1}$是上一次求估计时已计算过了，而$1+\varphi(\Phi^T\Phi)^{-1}\varphi^T$是一个数，这里求逆运算就成了一个数的除法。

(3-86)式具有很强的直观意义：

1. 第N+1次估计$\hat{\theta}(N+1)$，是第N次估计$\hat{\theta}(N)$加上一个修正项$K(N)[y_{N+1}-\varphi\hat{\theta}(N)]$

2. 因为

$$\varphi\hat{\theta}(N)=(-y_{n+N},-y_{n+N-1},\cdots,-y_{N+1},u_{n+N},u_{n+N-1},\cdots,u_{N+1})\hat{\theta}(N)$$

$$=-\hat{a}_1y_{n+N}-\hat{a}_2y_{n+N-1}-\cdots-\hat{a}_ny_{N+1}+\hat{b}_1u_{n+N}+\cdots+\hat{b}_nu_{N+1}$$

因此$\varphi\hat{\theta}(N)$表示估计得$\hat{\theta}(N)$后用(3-68)所得的予报值，y_{n+N+1}是新的实测值。

如果实测值与予报值相等，即

$$y_{n+N+1}=\varphi\hat{\theta}(N)$$

那么 $\hat{\theta}(N+1)=\hat{\theta}(N)$ 此时不必修改$\hat{\theta}(N)$

3. 如果 $y_{n+N+1}\neq\varphi\hat{\theta}(N)$，此时要得出$\hat{\theta}(N+1)$必须对$\hat{\theta}(N)$加以修改。修正项与$y_{N+n+1}-\varphi\hat{\theta}(N)$成正比，$K(N)$是增益因子，$y_{n+N+1}$与$\varphi\hat{\theta}(N)$的差愈大修正项愈大，增益因子愈大修正项愈大。

但是，我们必须指出(3-86)给出的递推公式是不完全的。为了得出K(N)的递推关系，我们定义

$$P(N)=\alpha[\Phi^T(N)\Phi(N)]^{-1} \tag{3-87}$$

当e_i是独立同分布随机变量时，式中α是e_i的方差

$$K(N)=(\Phi^T\Phi)^{-1}\varphi^T(1+\varphi(\Phi^T\Phi)^{-1}\varphi^T]^{-1}$$

$$=\alpha(\Phi^T\Phi)^{-1}\varphi^T[\alpha+\varphi\alpha(\Phi^T\Phi)^{-1}\varphi^T]^{-1}$$

由此得

$$K(N)=P(N)\varphi^T[\alpha+\varphi P(N)\varphi^T]^{-1} \tag{3-88}$$

(3-84)式两边乘α得

$$P(N+1)=P(N)-K(N)\varphi P(N)$$

$$=[I-K(N)\varphi]P(N) \tag{3-89}$$

把(3-86)，(3-88)，(3-89)集中起来，重新添上推导过程

中省略的N和N+1得到如下一组递推公式：

$$K(N)=P(N)\varphi^T(N+1)\left[\lambda+\varphi(N+1)P(N)\varphi^T(N+1)\right]^{-1} \tag{3-90}$$

$$P(N+1)=\left[I-K(N)\varphi(N+1)\right]P(N) \tag{3-91}$$

$$\hat{\theta}(N+1)=\hat{\theta}(N)+K(N)\left\{y_{n+N+1}-\varphi(N+1)\hat{\theta}(N)\right\} \tag{3-92}$$

利用(3-90)、(3-91)，(3-92)可以递推估计θ的值。

例2、用递推最小二乘法，在线辨识下列弹跳方程：

$$h=S_0+K_0F \tag{3-93}$$

其中h表示出口厚度，S_0是相对辊缝，K_0是机架刚度的倒数，F是总压力。

这种辨识可用于冷轧带钢轧制，用h的予报值来校正轧机的开口度，使带钢成品厚度一定，在这个问题中，S_0和K_0由于温升轧辊变形等影响，这两个参数是随着轧制过程的进行而变化的。因此，要求在线辨识这二个参数，才能使弹跳方程符合实际运行情况。

假设，我们通过仪表测得一组观测值$\{h_i,F_i\}$，可以写出其观测方程：

$$\begin{cases} h_1=S_0+K_0F_1+e_1 \\ h_2=S_0+K_0F_2+e_2 \\ \cdots \\ h_N=S_0+K_0F_N+e_N \end{cases}$$

如果令

$$\Phi(N)=\begin{pmatrix} 1 & F_1 \\ 1 & F_2 \\ \vdots & \\ 1 & F_N \end{pmatrix},\quad Y(N)=\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_N \end{pmatrix},\quad \theta=\begin{pmatrix} S_0 \\ K_0 \end{pmatrix} \tag{3-94}$$

设e_i是独立同分布随机变量（量测误差）且e_i的方差等于1。对于这一稳态方程的辨识，仍可以运用(3-90)—(3-92)公式组。其中求和号下标起始值略有不同。

把(3-90)代入(3-91)得

$$P(N+1)=P(N)-P(N)\varphi^T(N+1)[1+\varphi(N+1)P(N)\varphi^T(N+1)]^{-1}\varphi(N+1)P(N)$$

(3－95)

又(3－91)可写成

$$P(N+1)P^{-1}(N)=I-K(N)\varphi(N+1)$$

$$K(N)\varphi(N+1)=I-P(N+1)P^{-1}(N)$$

$$=P(N+1)[P^{-1}(N+1)-P^{-1}(N)]$$

因为 $P^{-1}(N+1)-P^{-1}(N)=\varphi^T(N+1)\varphi(N+1)$

得 $K(N)\varphi(N+1)=P(N+1)\varphi^T(N+1)\varphi(N+1)$

由此可知

$$K(N)=P(N+1)\varphi^T(N+1) \qquad (3-96)$$

把(3－96)代入(3－92)得

$$\hat{\theta}(N+1)=\hat{\theta}(N)+P(N+1)\varphi^T(N+1)[y_{N+1}-\varphi(N+1)\hat{\theta}(N)]$$

(3－97)

这样(3－95)和(3－97)组成了一对递推公式：

$$\boxed{\begin{aligned}&P(N+1)=P(N)-P(N)\varphi^T(N+1)[1+\varphi(N+1)P(N)\varphi^T(N+1)]^{-1}\varphi(N+1)P(N) \quad (3-95)\\&\hat{\theta}(N+1)=\hat{\theta}(N)+P(N+1)\varphi^T(N+1)[y_{N+1}-\varphi(N+1)\hat{\theta}(N)] \quad (3-97)\\&P(N)=(\Phi^T(N)\Phi(N))^{-1}\end{aligned}}$$

在本例中采用这一组递推公式。

设

$$P(N)=[\Phi^T(N)\Phi(N)]^{-1}=\begin{pmatrix}p_{11}(N) & p_{12}(N)\\ p_{21}(N) & p_{22}(N)\end{pmatrix}$$

而 $\varphi(N+1)=(1,F_{N+1})$

设

$$P(N+1)=\begin{pmatrix}p_{11}(N+1) & p_{12}(N+1)\\ p_{21}(N+1) & p_{22}(N+1)\end{pmatrix}$$

根据(3－95)，因为

$$P(N)\varphi^{T}(N+1)=\begin{pmatrix} p_{11}(N) & p_{12}(N) \\ p_{21}(N) & p_{22}(N) \end{pmatrix}\begin{pmatrix} 1 \\ F_{N+1} \end{pmatrix}$$

$$=\begin{pmatrix} p_{11}(N)+F_{N+1}\,p_{12}(N) \\ p_{21}(N)+F_{N+1}\,p_{22}(N) \end{pmatrix}$$

$$\varphi(N+1)P(N)\varphi^{T}(N+1)$$

$$=p_{11}(N)+F_{N+1}\,p_{21}(N)+F_{N+1}\,p_{12}(N)+F_{N+1}^{2}\,p_{22}(N)$$

$$\varphi(N+1)P(N)=(p_{11}(N)+F_{N+1}\,p_{21}(N),\ p_{12}(N)+F_{N+1}\,p_{22}(N))$$

所以有

$$\begin{cases} p_{11}(N+1)=p_{11}(N)-(p_{11}(N)+p_{12}(N)F_{N+1})^{2}[p_{11}(N)+F_{N+1}\,p_{12}(N \\ \qquad +(p_{12}(N)+p_{22}(N)F_{N+1})F_{N+1}+1]^{-1}; \\ p_{12}(N+1)=p_{21}(N+1)=p_{12}(N)-(p_{11}(N)+p_{12}(N)F_{N+1})(p_{12}(N) \\ \qquad +p_{22}(N)F_{N+1})\cdot[p_{11}(N)+p_{12}(N)F_{N+1}+(p_{12}(N) \\ \qquad +p_{22}(N)F_{N+1})F_{N+1}+1]^{-1}; \\ p_{22}(N+1)=p_{22}(N)-(p_{12}(N)+p_{22}(N)F_{N+1})^{2}[p_{11}(N)+p_{12}(N)F_{N+1} \\ \qquad +(p_{12}(N)+p_{22}(N)F_{N+1})F_{N+1}+1]^{-1} \end{cases} \tag{3-98}$$

(3-98)这一组公式给出了协方差矩阵每个元素的递推关系式、

又
$$\hat{\theta}_{n}=\begin{pmatrix} \hat{S}_{0}(N) \\ \hat{K}_{0}(N) \end{pmatrix},\quad y_{N+1}=h_{N+1}$$

因此可得

$$y_{N+1}-\varphi(N+1)\hat{\theta}(N)=h_{N+1}-\hat{S}_{0}(N)-\hat{K}_{0}(N)F_{N+1} \tag{3-99}$$

把(3-99)代入(3-97)得

$$\hat{S}_{0}(N+1)=\hat{S}_{0}(N)+(p_{11}(N+1)+p_{12}(N+1)F_{N+1})$$

$$(h_{N+1}-\hat{S}_{0}(N)-\hat{K}_{0}(N)F_{N+1}) \tag{3-100}$$

$$\hat{K}_{0}(N+1)=\hat{K}_{0}(N)+(p_{12}(N+1)+p_{22}(N+1)F_{N+1})(h_{N+1}-\hat{S}_{0}(N)-\hat{K}_{0}(N)F_{N+1})$$

(3-100)

我们用(3−98)，(3−100)计标出各个时刻的$\hat{S}_0(N)$和$\hat{K}_0(N)$，并用下式来予报轧出厚度，予报值记作$\hat{h}_{N+1}$

$$\hat{h}_{N+1} = \hat{S}_0(N) + \hat{K}_0(N) F_{N+1}$$

这样我们可以用这次观测得的F_{N+1}，与上一次估计得的$S_0(N)$，$K_0(N)$予测h_{N+1}之值。

利用(3−90)，(3−91)，(3−92)一组递推公式或(3−95)，(3−97)递推公式都涉及到如何选取$P(N)$和$\hat{\theta}(N)$初值的问题。

这里提供二种方法：

1. 先取一批观测值，用最小二乘法估计出θ，作为初值。此时必须注意下面的情况：

设

$$\Phi(N) = \begin{bmatrix} \varphi(1) \\ \varphi(2) \\ \vdots \\ \varphi(N) \end{bmatrix}$$

那么 $$\Phi^T(N)\Phi(N) = \sum_{t=1}^{N} \varphi^T(t)\varphi(t) \qquad (3-101)$$

如果观测个数N小于未知参数的个数，在(3−68)模型中，即

$$N < 2n$$

此时由观测方程介出的$a_1, \cdots, a_n, b_1, \cdots, b_n$不唯一，$\Phi^T\Phi$不是正定矩阵，为了找到正定的协方差初值，我们必须取$N_0 > 2n$（参数个数）。

用 $$P_0 = P(N_0) = \alpha\left[\Phi^T(N_0)\Phi(N_0)\right]^{-1} \qquad (3-102)$$

$$\hat{\theta}_0 = \left[\Phi^T(N_0)\Phi(N_0)\right]^{-1}\Phi^T(N_0)y(N_0) \qquad (3-103)$$

作为初值。

2. 人为地给定初值：

$$\left.\begin{aligned} P_0 &= \frac{1}{\varepsilon} I \quad \varepsilon\text{足够小} \\ \hat{\theta}_0 &= 0 \end{aligned}\right\} \qquad (3-104)$$

可以证明采用这一组初值，经过一定次数的递推计祘就能得到满意的结果，详细证明见[6]。

§3.3.3　关于系统阶数递推的参数最小二乘估计

在实际辨识问题中，当测得了一系列观测值$\{u(t), y(t)\}$ $t=1,2,\cdots,n+N$时，我们并不知道系统的确切阶数。因此，为了确定系统的阶，往往需要对各种阶数的参数分别进行最小二乘估计。如何利用低阶系统模型的参数估计值来递推计祘高阶系统模型的参数估计值，这就是关于系统阶数递推的参数最小二乘估计所要解决的问题。

为了推导书写的方便，我们不妨设系统模型采用低阶数时包含K个参数，采用高阶数模型时包含n个参数，$n>K$。把Φ矩阵写成分块矩阵的形，Φ中的元素用φ_{ij}来表示那么有

$$\Phi_n=\left(\begin{array}{cccc|cccc}\varphi_{11} & \varphi_{12} & \cdots & \varphi_{1K} & \varphi_{1,K+1} & \cdots & & \varphi_{1n}\\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2K} & \varphi_{2,K+1} & \cdots & & \varphi_{2n}\\ \vdots & & & \vdots & \vdots & & & \vdots\\ \varphi_{N1} & \varphi_{N2} & \cdots & \varphi_{NK} & \varphi_{N,n+1} & & & \varphi_{Nn}\end{array}\right)=(\phi_1 \vdots \phi_2) \tag{3-105}$$

把参数向量写成相应的分块形式

$$\theta=\begin{pmatrix}\theta_1\\ \vdots\\ \theta_K\\ \hline \theta_{K+1}\\ \vdots\\ \theta_n\end{pmatrix}=\begin{pmatrix}\Theta_1\\ \hline \Theta_2\end{pmatrix} \tag{3-106}$$

其中Θ_1即低阶模型所含的未知参数，Θ_2表示用高阶模型时增加的参数。

此时最小二乘法所得的正规方程

$$(\Phi_n^T\Phi_n)\theta=\Phi_n^T Y \tag{3-107}$$