

# 小样本的线性连续时间系统的辨识

秦霞 韩建勋

(天津商学院) (天津大学)

## 摘要

本文简要地讨论了各种辨识线性连续时间系统的方法,并根据 *J. Aitchison* 提出的预报拟合优度准则<sup>[5]</sup>,建立了一种新的连续时间系统在小样本条件下的辨识方法.模拟仿真实验的结果表明:当量测数据很少且噪声较大时,该方法明显优于其它方法.

## 一、引言

在生物医学及化工等领域研究数学模型时,经常遇到的问题是:所需的必要信息一量测数据非常少,用以往的传统方法来估计数学模型的参数时,精度往往很低,不能满足要求.在这种情况下,如何进一步提高待估参数的精度,即小样本条件下的系统辨识问题,成为人们普遍关心的一个问题.本文就这个问题初步做了一些探讨.模拟实验表明:文中所给出的方法,对于连续系统的小样本参数估计,其精度比传统方法有所提高,并且具有其它方法所没有的独特优点.

## 二、数学模型的描述

生物医学领域中,常见的机理模型是房室模型<sup>[9]</sup>,用数学上的术语来描述,即其数学模型的类型为线性连续系统的  $n$  阶常微分方程组,化工过程的数学模型也有一些与其类似.

假设  $n$  阶线性连续系统的数学模型为如下形式<sup>[8]</sup>:

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}, \quad (1)$$

$$\mathbf{y} = C\mathbf{x}. \quad (2)$$

其中  $\mathbf{x}$  为  $n$  维状态向量,  $\mathbf{u}$  为  $r$  维输入向量,  $\mathbf{y}$  为  $q$  维输出向量,  $A$ 、 $B$ 、 $C$  分别为  $n \times n$  阶系统矩阵、 $n \times r$  阶输入矩阵、 $q \times r$  阶输出矩阵.

式(1)和式(2)是线性连续系统模型的一般形式,它同时也是生物医学系统房室模型的数学模型形式.在具体应用时,模型可以有比较简单的形式.

### 三、辨识方法种种

连续系统模型的辨识,大致可分为两类。一类是间接辨识方法,即先将连续系统的微分方程离散化,变为差分方程,运用离散系统辨识的各种方法之一,估计离散系统的参数。然后根据离散模型与连续模型的对应关系,最终建立连续系统的模型;另一类方法是直接辨识方法,目前已有几种可用的方法,例如 Walsh 函数法<sup>[3]</sup>,块脉冲函数法<sup>[2]</sup>,最近又出现了利用数值积分的梯形法则的代数学方法<sup>[2,6]</sup>,以及样条函数辨识方法<sup>[4]</sup>。以上辨识方法对于大样本量测是比较可行的,即只要样本足够大,便可以得到比较满意的辨识精度。但是,对于因受到种种客观条件的限制而造成的小样本量测,用上述方法辨识往往达不到满意的效果。

在生物医学及其它领域,目前普遍应用的小样本辨识方法有两种。一种方法是在采样设计上运用最优采样方案<sup>[1]</sup>,它是依据把 Fisher 信息矩阵的逆做为参数估计协方差矩阵的下限这个准则而设计最优采样的,文献 [1] 中以实例证明了这种方法对于仅限于小样本离散时间上的输出量测是非常有效的;另一种方法是非线性最小二乘法<sup>[7]</sup>,这种方法就是先对任何给定的参数值根据初始条件解微分方程模型,得到状态的时间函数,然后再计算出模型的输出,最后将采样时间点上的模型输出与实验数据相比较,按照非线性最小二乘准则来确定参数。上述两种方法对于小样本辨识具有一定的有效性,但也存在一些缺陷,如最优采样的方法只能适用于某些有条件采取这种方案的场合,而且在理论设计上,这种方法也比较复杂;非线性最小二乘法由于需要求解模型,在模型形式较复杂时,应用起来很不方便。因此,有必要寻求一种既简单又有效的连续系统的小样本辨识方法。

### 四、小样本极大似然法

小样本极大似然法 (Small Sample Maximum Likelihood Method, 简称为 SSML 方法)是本文提出的一种新的小样本辨识方法。这种方法的理论依据是 J. Aitchison 提出的预报拟合优度准则及其在多变量正态情形中的应用的结论<sup>[5]</sup>。SSML 方法概要如下:通常,对于属于正态总体的数据的概率密度拟合是使用正态分布,这在大样本的情况下,拟合是较好的。但是在小样本的情况下,这种拟合存在较大偏差,样本越小,偏差越大。J. Aitchison 提出的预报拟合优度准则解决了这个问题。他运用这个准则,在多变量正态情形中得出如下结论:即对于正态总体的小样本数据,其最优的概率密度拟合是学生 (Student) 分布。因此,依据极大似然定理,可推出小样本情形下,动态系统模型参数的极大似然估计。

根据 J. Aitchison 的结论,对于正态总体的小样本数据来说, $d$  维随机向量  $\mathbf{y}$  的概率密度拟合为

$$p(\mathbf{y}|\boldsymbol{\theta}) = \text{Std}(\nu, \mathbf{m}, \Sigma),$$

即

$$p(\mathbf{y}|\boldsymbol{\theta}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\pi^{\frac{d}{2}}|\nu\Sigma|^{\frac{1}{2}}\Gamma\left(\frac{\nu-d+1}{2}\right)} \{1 + (\mathbf{y} - \mathbf{m})^T(\nu\Sigma)^{-1}(\mathbf{y} - \mathbf{m})\}^{-\frac{\nu+1}{2}}. \quad (3)$$

其中,  $\nu = N - 1$ ,  $\nu$  为自由度,  $N$  为样本数,  $\mathbf{m}$  为  $\mathbf{y}$  的均值向量,  $\Sigma$  为  $\mathbf{y}$  的协方差矩阵, Std 表示学生分布.

由 (3) 式出发, 应用极大似然定理, 便可推导出小样本极大似然法的计算公式.

考虑  $d$  维随机向量  $\mathbf{y}$  的  $N$  个独立的观测  $\{\mathbf{y}_t, t = 1, \dots, N\}$ , 每一个  $\mathbf{y}_t$  具有均值  $\mathbf{x}_t\boldsymbol{\theta}$  及相同协方差矩阵  $\Sigma$  的学生分布, 则  $\boldsymbol{\theta}$  和  $\Sigma$  是极大似然估计量的必要条件为

$$\boldsymbol{\theta} = \left[ \sum_{t=1}^N \frac{\mathbf{x}_t^T(\nu\Sigma)^{-1}\mathbf{x}_t}{1 + (\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta})^T(\nu\Sigma)^{-1}(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta})} \right]^{-1} \cdot \left[ \sum_{t=1}^N \frac{\mathbf{x}_t^T(\nu\Sigma)^{-1}\mathbf{y}_t}{1 + (\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta})^T(\nu\Sigma)^{-1}(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta})} \right], \quad (4)$$

$$\Sigma = \left[ \frac{\nu+1}{N\nu} \sum_{t=1}^N \frac{(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta})^T}{1 + (\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta})^T(\nu\Sigma)^{-1}(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta})} \right]^T. \quad (5)$$

式 (4) 和式 (5) 即为本文推导的最终结果, 其中  $\boldsymbol{\theta}$  为待估参数向量. 由于  $\boldsymbol{\theta}$  和  $\Sigma$  为隐式形式, 不易直接求解, 故采用了如下的松弛算法:

(i)  $\boldsymbol{\theta}$ 、 $\Sigma$  取任一值, 如  $\boldsymbol{\theta}$  取最小二乘估计值  $\boldsymbol{\theta}_{LS} = (\mathbf{x}_t^T\mathbf{x}_t)^{-1}\mathbf{x}_t^T\mathbf{y}_t$ ,  $\Sigma$  取单位矩阵  $I$ , 且设初值为  $\boldsymbol{\theta}_0$ 、 $\Sigma_0$ .

(ii) 将式 (4) 和式 (5) 改写为迭代形式如下:

$$\boldsymbol{\theta}_{k+1} = \left[ \sum_{t=1}^N \frac{\mathbf{x}_t^T(\nu\Sigma_k)^{-1}\mathbf{x}_t}{1 + (\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta}_k)^T(\nu\Sigma_k)^{-1}(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta}_k)} \right]^{-1} \times \left[ \sum_{t=1}^N \frac{\mathbf{x}_t^T(\nu\Sigma_k)^{-1}\mathbf{y}_t}{1 + (\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta}_k)^T(\nu\Sigma_k)^{-1}(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta}_k)} \right], \quad (6)$$

$$\Sigma_{k+1} = \left[ \frac{\nu+1}{N\nu} \sum_{t=1}^N \frac{(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta}_k)(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta}_k)^T}{1 + (\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta}_k)^T(\nu\Sigma_k)^{-1}(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\theta}_k)} \right]^T. \quad (7)$$

其中,  $k$  为迭代次数,  $k = 0, 1, \dots$ .

利用前一次的迭代结果  $\boldsymbol{\theta}_k$ 、 $\Sigma_k$  代入式 (6) 及式 (7) 中, 可求出新的迭代结果  $\boldsymbol{\theta}_{k+1}$ 、 $\Sigma_{k+1}$ .

(iii) 若  $|\boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k| < \varepsilon_1$ ,  $|\Sigma_{k+1} - \Sigma_k| < \varepsilon_2$  (其中  $\varepsilon_1$ 、 $\varepsilon_2$  为预先给定的迭代精度), 则停止计算, 且取  $\boldsymbol{\theta} = \boldsymbol{\theta}_{k+1}$ ,  $\Sigma = \Sigma_{k+1}$ ; 否则转向 (ii).

线性预报误差模型

$$\mathbf{y}_t = \mathbf{x}_t\boldsymbol{\theta} + \varepsilon_t \quad (8)$$

可直接应用 SSML 方法, 而连续系统的微分方程模型不能直接利用 SSML 方法, 必须先变为上述形式的预报误差模型. 对于本文第二节给出的模型形式, 即式 (1) 和式 (2), 可先将其化为输入—输出模型如下.

对式 (1) 和式 (2) 的两边取拉氏变换, 得

$$SX(S) = AX(S) + BU(S),$$

$$Y(S) = CX(S),$$

则有

$$Y(S) = C(SI - A)^{-1}BU(S). \quad (9)$$

上式又可改写为

$$Y(S) = C \cdot \frac{\text{adj}(SI - A)}{\det(SI - A)} \cdot B \cdot U(S). \quad (10)$$

由此可知,它是一个  $n$  阶输入—输出系统,故式 (10) 可化为如下的连续系统微分方程模型形式:

$$\begin{aligned} \frac{d^n y}{dt^n} + P_1(\theta) \frac{d^{n-1} y}{dt^{n-1}} + \dots + P_n(\theta) y &= Q_0(\theta) \frac{d^m u}{dt^m} \\ + Q_1(\theta) \frac{d^{m-1} u}{dt^{m-1}} + \dots + Q_m(\theta) u. \end{aligned} \quad (11)$$

其中  $m$  为式 (10) 分子的阶数,  $n$  为分母阶数即系统阶次,  $\theta$  为未知参数按一定顺序构成的向量,  $P_i(\theta) (i = 1, \dots, n)$  及  $Q_j(\theta) (j = 1, \dots, m)$  为与  $\theta$  有关的参数矩阵. 因此,估计未知模型的参数,可利用式 (11) 的输入—输出模型.

下面的问题就是如何将式 (11) 变成式 (8) 形式的预报误差模型. 我们参考 M. S. Ahmed 提出的辨识线性连续时间系统的样条函数法<sup>[4]</sup>来简化模型 (11) 式, 便得到可直接应用 SSML 方法的预报误差模型.

为了评价 SSML 方法,我们采用文献 [4] 中所提供的关于标准化误差的定义做为判别模拟仿真实验精度的准则. 标准化误差 (N. E.) 定义为

$$N. E. = \frac{\|\hat{\theta} - \theta\|^2}{\|\theta\|^2}, \quad (12)$$

$\hat{\theta}$  为参数估计值,  $\theta$  为参数真值. 因篇幅所限,本文仅给出一阶及二阶模型各一例的模拟仿真实验结果,见表 1 及表 2. 我们采用三种辨识方法进行了对比实验,即非线性最小二乘法,以 NLLS 表示;均匀采样的小样本极大似然法,以 SSML 表示;最优采样的小样本极大似然法,以 OSSML 表示.

表 1 一阶模型仿真实验结果

模 型		$\dot{x} + ax = bu,$		$y_k = x_k + \delta_k$	
参数真值		$a = 1, \quad b = 1$		初 值	$x(0) = 3$
输 入		$u = 1(t)$		噪声方差	$\sigma_\delta^2 = 0.1$
样 本 数		6	8	10	12
N. E.	NLLS	$2.752 \times 10^{-1}$	$1.725 \times 10^{-1}$	$1.206 \times 10^{-1}$	$9.898 \times 10^{-2}$
	SSML	$1.112 \times 10^{-1}$	$6.376 \times 10^{-2}$	$5.277 \times 10^{-2}$	$4.965 \times 10^{-2}$
	OSSML	$5.580 \times 10^{-2}$	$3.376 \times 10^{-2}$	$1.297 \times 10^{-2}$	$8.584 \times 10^{-3}$

将表 1、表 2 中数据绘制成曲线,如图 1 及图 2 所示.

由上述模拟实验的表格数据及曲线图可以看出:小样本极大似然法的参数估计精度比非线性最小二乘法有所提高,样本越小,效果越好.此外,最优采样比均匀采样的效

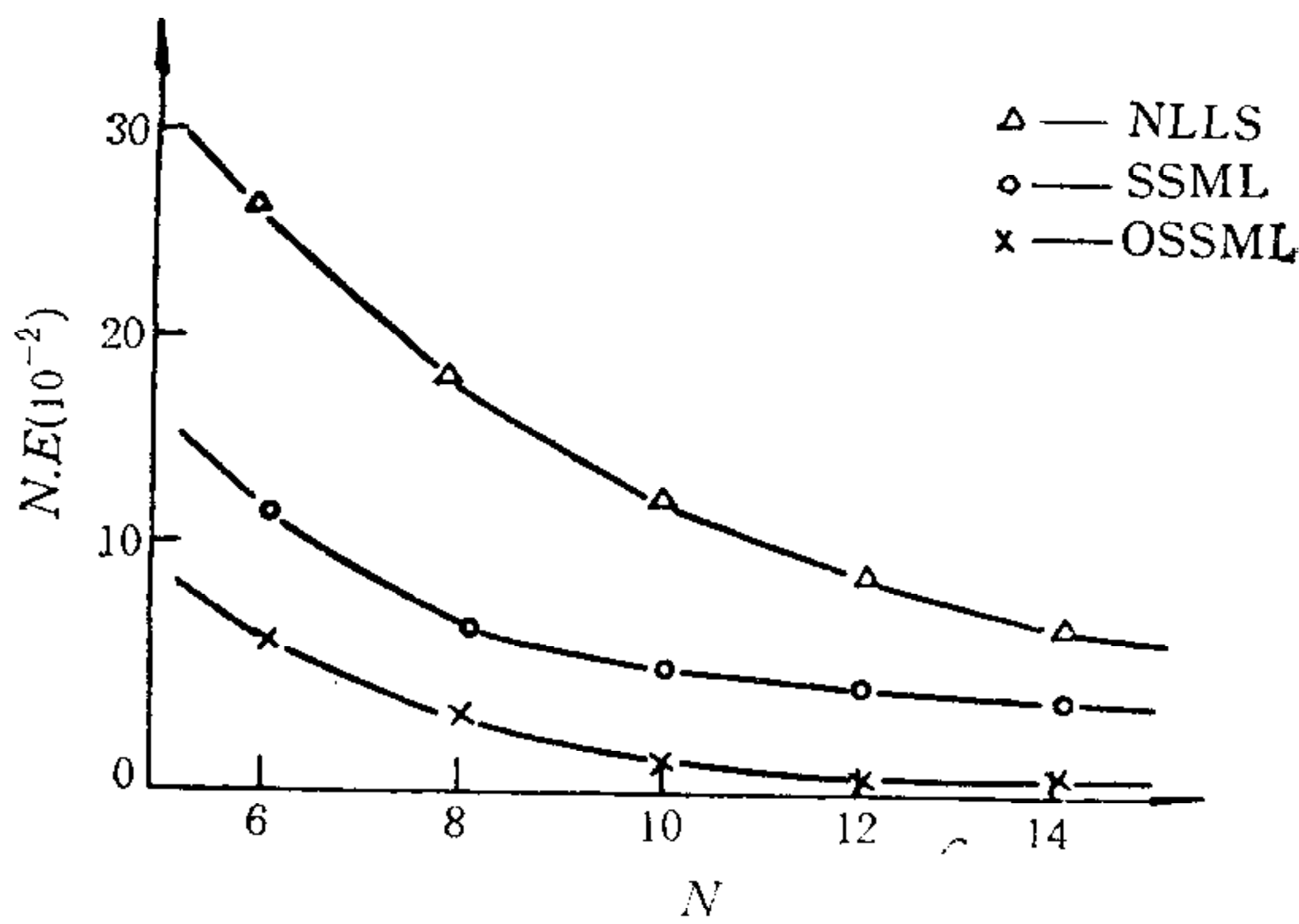


图 1 一阶模型 N-N. E. 曲线

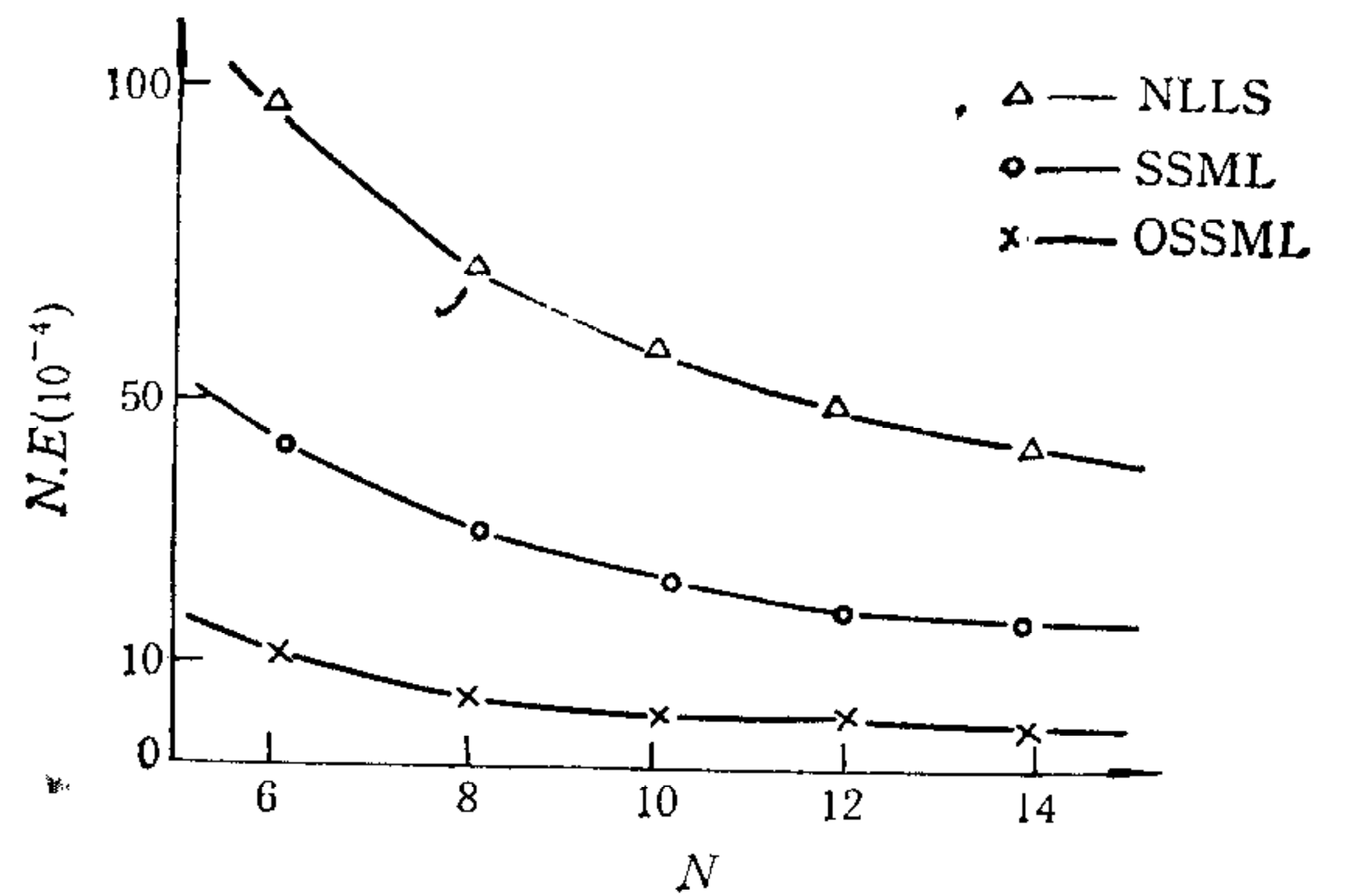


图 2 二阶模型 N-N. E. 曲线

表 2 二阶模型仿真实验结果

模 型	$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} u, y = (01) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \varepsilon$				
参数真值	$a_{11} = -1.5, a_{12} = 0.4, a_{21} = -0.8, a_{22} = 0.2, b = -0.5$				
初 值	$x_1(0) = x_2(0) = 1.2$	输 入	$u = 1(t)$		
噪声方差	$\sigma_\varepsilon^2 = 0.1$				
样 本 数	6	8	10	12	
N. E.	NLLS	$9.968 \times 10^{-3}$	$7.278 \times 10^{-3}$	$6.266 \times 10^{-3}$	$5.097 \times 10^{-3}$
	SSML	$4.459 \times 10^{-3}$	$3.384 \times 10^{-3}$	$2.823 \times 10^{-3}$	$1.856 \times 10^{-3}$
	OSSML	$9.282 \times 10^{-4}$	$7.14 \times 10^{-4}$	$6.129 \times 10^{-4}$	$5.088 \times 10^{-4}$

果更佳。样本数增大，三种方法的 N-N. E. 曲线趋于平稳。因此，可以说 SSML 方法是解决小样本问题的比较有效的方法。

### 五、实 例

现有一肾衰竭病人在人工肾透析过程中，尿素氮的一组浓度随时间变化的采样数据：

$t_i$ (小时)	0	1	2	3	4	5	6
$y_i$ (mg/dl)	96	75	62	52	45	41	37

由透析机理，可写出如下物理模型：

$$V \frac{dc(t)}{dt} = -C_L \cdot c(t) + G, \tag{5.1}$$

其中， $V$  为人体体液量 (dl)， $c(t)$  为人体内尿素氮瞬时浓度 (mg/dl)， $C_L$  为尿素氮清除率 (dl/h)， $G$  为人体内尿素氮产生速率 (mg/h)。

令  $c(t) = x(t)$ ， $-C_L/V = a$ ， $G/V = b$ ，则 (5.1) 式可化为以  $x(t)$  为状态变量的数学模型：

$$\frac{dx(t)}{dt} = ax(t) + b. \quad (5.2)$$

设量测量为  $y(t_i) = y_i (i = 0, 1, \dots, 6)$ , 量测噪声为  $\varepsilon_i$ , 则有量测方程为

$$y(t_i) = x(t_i) + \varepsilon_i \quad (5.3)$$

式 (5.2) 和 (5.3) 即为一阶状态空间模型.

我们用前六个量测数据进行辨识, 得出小样本极大似然法的参数估计结果:  $a = -0.367$ ,  $b = 11.049$ . 因此透析过程的数学模型可写为

$$\frac{dx(t)}{dt} = -0.367x(t) + 11.049. \quad (5.4)$$

由上式得模型的解析解为

$$x(t) = 30.106 + 65.894e^{-0.367t}. \quad (5.5)$$

令  $t = 6(h)$ , 则得  $x(6) = 37.393(mg/dl)$ . 因为量测  $y(6) = y_6 = 37(mg/dl)$ , 其相对误差为 1%.

## 六、结 论

通过前面的理论分析及模拟实验结果, 可以看到 SSML 方法具有如下特点:

(1) 计算比较简单. 因为它是利用松弛算法进行迭代计算的, 编程容易, 计算时间也较短.

(2) 对小样本条件下的参数估计较为有效和适用.

(3) 在适用范围方面, 可以进一步推广应用. SSML 方法是以线性预报误差模型作为基础模型的, 故其它形式的线性系统模型, 凡是能直接或间接地转化成预报误差模型的, 一般均可使用 SSML 方法.

(4) 在应用条件上, SSML 方法从理论上讲, 仅限于正态噪声干扰的小样本量测数据; 而对于非正态干扰及未知噪声特性的样本数据, 还需作进一步的探讨和研究.

综上所述, 可以认为 SSML 方法是值得推广应用的一种小样本辨识方法.

## 参 考 文 献

- [1] Fumihika Mori and Distefano, J. J., *IEEE Trans.* **AC-24**(1979), 893—900.
- [2] Sinha, N. K., Lastman, G. J., *Int. J. Control*, **35**(1962), 117—126.
- [3] Bohn, E. V., *IEEE Trans.* **IE-29**(1982), 38—46.
- [4] Ahmed, M. S., *Modeling and Simulation*, Vol. **13**, pt. 1(1982), 171—177.
- [5] Aitchison, J., *Biometrika*, **62**(1975), 547—554.
- [6] Hung J. C., Liu C. C., Chan P. Y., *Proc. 14th Asilomar Conference on Circuits, Systems and Computers*, 24—34, 1980.
- [7] A. Van den bos, *Automatica*, **16**(1980), 487—490.
- [8] 哥德温, G. C., 潘恩, R. L., *动态系统辨识——试验设计与数据分析*, 科学出版社 (1983).
- [9] 峰島三千男, 酒井清孝, *計測と制御*, **19**(1980), 84—87.

---

# THE IDENTIFICATION OF LINEAR CONTINUOUS-TIME SYSTEM FOR SMALL SAMPLES OF INPUT- OUTPUT DATA

QIN XIA

*(Tianjin Commercial College)*

HAN JIANXUN

*(Tianjin University)*

## ABSTRACT

In this paper, different methods for the identification of linear continuous-time systems are concisely discussed. Moreover, according to the criterion on goodness of prediction fit proposed by J. Aitchison<sup>[5]</sup>, a new method for identifying continuous-time system limited to small samples is established. The results from simulation experiments indicate that when the data obtained are very few and quite noisy, this method is obviously superior to others.