

非线性系统的多模态 ARMAX 模型及其参数辨识

周超俊 蒋慰孙

藤井省三

(华东理工大学自动化研究所·上海, 200237) (名古屋大学情报工学科·日本)

摘要: 本文利用非线性系统多个工作点上的线性化模型, 建立了一类参数内插型多模态 ARMAX 模型. 与传统的线性化模型相比较, 在更大的范围内有较高的精度. 本文并给出了如何辨识多模态模型参数的方法. 仿真例子表明了多模态模型及其参数辨识算法的有效性.

关键词: 非线性系统; 参数辨识; 多模态模型

1 引言

非线性系统的建模是一个相当困难的问题. 基于机理所得到的模型往往过于复杂而难以使用. 传统的线性化模型则难以在大范围内得到较高的精度. 近年来出现的基于神经网络的建模方法虽较有效, 但也存在局部极小值和建模时计算量相当大的问题.

传统的基于某一点的线性化模型固然难以在大范围内有较高的精度, 然而如能利用多个工作点上的线性化模型来近似非线性系统, 却可望获得较高的精度. 这正是本文所进行的主要工作.

2 非线性系统的参数内插型多模态 ARMAX 模型

考虑图 1 所示的闭环 SISO 非线性系统. 图中 C 是控制器, r 是系统的设定值, G_0 是非线性对象, 其状态表述式为

$$\dot{x} = f(x, u), \quad (1)$$

$$y = g(x, u). \quad (2)$$

式中 x , u 和 y 分别是状态向量(n 维), 输入变量(l 维) 和输出变量(l 维).

假设在稳态时闭环系统无余差, 则有

$$e(\infty) = r - y(\infty) = 0. \quad (3)$$

对应于给定的设定值 r , 可以求得该非线性系统的一个平衡点 (x^*, u^*) , 利用泰勒展开方法, 可得到该系统在工作点 (x^*, u^*) 的线性化模型

$$\Delta \dot{x} = \left. \frac{\partial f}{\partial x} \right|_{\substack{x=x^* \\ u=u^*}} \Delta x + \left. \frac{\partial f}{\partial u} \right|_{\substack{x=x^* \\ u=u^*}} \Delta u, \quad (4)$$

$$\Delta y = \left. \frac{\partial g}{\partial x} \right|_{\substack{x=x^* \\ u=u^*}} \Delta x + \left. \frac{\partial g}{\partial u} \right|_{\substack{x=x^* \\ u=u^*}} \Delta u. \quad (5)$$

然而这样的做法会丢失大量关于原来非线性系统的静态和动态信息. 因此, 当设定值 r 有较大变时, 线性模型的精度难以保证. 既然上述线性化模型的精度与工作点有关, 亦即与

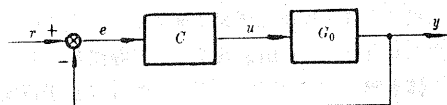


图 1 闭环非线性系统

设定值 r 有关,那么如能把 r 的信息引入线性化模型,所得的新模型将具有较高的精度.从这样的考虑出发,设令设定值 r 的变化范围为 $r \in [r_{\min}, r_{\max}]$,在该区间中取 N 个离散的点,则可得到闭环系统的 N 个可能的设定点,分别记为

$$w_r = \{r^i, i = 1, 2, \dots, n\}. \quad (6)$$

对应的工作点是

$$w_p = \{(x^1, u^1), (x^2, u^2), \dots, (x^n, u^n)\}. \quad (7)$$

在这一组工作点上,分别可求取象式(4)和(5)形式的线性化模型.同时,也不难转化为 ARMAX 模型

$$A^i(q^{-1})\Delta y_k^i = B^i(q^{-1})\Delta u_k^i, \quad (8)$$

式中的 q^{-1} 是后向算子, $\Delta y_k^i = y_k - r^i$, $\Delta u_k^i = u_k - u^i$, k 为采样次数, $A^i(q^{-1})$ 和 $B^i(q^{-1})$ 都是 q^{-1} 的多项式.

定义下列模型为参数内插型多模态 ARMAX 模型.

$$\sum_{i=1}^N \alpha_i(r) A^i(q^{-1}) \Delta y_k^i = \sum_{i=1}^N \alpha_i(r) B^i(q^{-1}) \Delta u_k^i. \quad (9)$$

式中的 $\alpha_i(r)$ 是模型参数,称为内插因子,它们各是设定值 r 的函数,满足下列条件

$$\sum_{i=1}^N \alpha_i(r) = 1, \quad \alpha_i(r^i) \doteq 1, \quad \alpha_i(r^j) \doteq 0, \quad i \neq j. \quad (10)$$

可以采用分段线性内插形式来计算内插因子.我们称式(8)为多模态模型的一个模态, N 为多模态模型的模态数.

由条件式(10)可以看到:在任一规定的工作点上,多模态模型就是非线性系统在该工作点上的线性化模型;而在规定工作点之间,多模态模型是由不同模态的参数内插得出的一个线性模型.尽管多模态模型式(9)对于 y_k 和 u_k 都是线性的,但模型参数随设定值而变化,从而保留着非线性变化的特征,这就使得式(9)有可能在更大的范围内有较高的精度.式(9)综合了非线性系统多个工作点上的信息,也可以看成是传统的线性化模型的一种自然推广.对于大多数工业过程来说,系统实际的工作点个数不多,因而多模态模型的参数个数也是十分有限的,这是一个有用的性质.

多模态模型的参数 $\alpha_i(r)$ 可以通过文[1,2]给出的方法近似计算,也可以用最小二乘的方法加以辨识,通过对辨识算法的分析,可以得到关于 $\alpha_i(r)$ 的一种物理意义,即权值的概念.我们把求取 $\alpha_i(r)$ 的问题称为多模态模型的参数辨识.

3 参数内插型多模态 ARMAX 模型的参数辨识

考虑测量噪声的存在,将式(9)改写为

$$\sum_{i=1}^N \alpha_i(r) [A^i(q^{-1})\Delta y_k^i - B^i(q^{-1})\Delta u_k^i] = e_k. \quad (11)$$

式中 e_k 为零均值的白噪声,

$$A^i(q^{-1}) = 1 + a_1^i q^{-1} + \dots + a_n^i q^{-n},$$

$$B^i(q^{-1}) = b_0^i + b_1^i q^{-1} + \dots + b_n^i q^{-n}.$$

定义误差向量 E_k

$$E_k = [e^1, e^2, \dots, e^N]^T. \quad (12)$$

式中

$$e^i = A^i(q^{-1})\Delta y_k^i - B^i(q^{-1})\Delta u_k^i. \quad (13)$$

则式(11)也可改写成矩阵方程形式

$$E_k^T \alpha_k = \varepsilon_k. \quad (14)$$

式中

$$\alpha_k = [\alpha_1(r), \alpha_2(r), \dots, \alpha_N(r)]^T, \quad (15)$$

α_k 是参数向量.

取目标函数为

$$J = \frac{1}{2} \sum_{k=1}^L \varepsilon_k^2 = \frac{1}{2} \sum_{k=1}^L (E_k^T \alpha_k)^2. \quad (16)$$

式中 L 为总的采样次数.

约束条件为

$$\alpha_k^T \mathbf{1}_N = 1, \quad (17)$$

式中

$$\mathbf{1}_N = [1, 1, \dots, 1]^T.$$

为了求解这个带约束条件的最小二乘问题,建立如下的拉格朗日函数

$$J = \sum_{k=1}^L \frac{1}{2} (E_k^T \alpha_k)^2 + \lambda_k (\alpha_k^T \mathbf{1}_N - 1). \quad (18)$$

求偏导后得到

$$\frac{\partial J}{\partial \alpha_k} = \sum_{k=1}^L E_k E_k^T \alpha_k + \lambda_k = 0, \quad (19)$$

$$\frac{\partial J}{\partial \lambda_k} = \mathbf{1}_N^T \alpha_k - 1 = 0. \quad (20)$$

写成矩阵形式可得

$$\begin{bmatrix} \sum_{k=1}^L E_k E_k^T & \mathbf{1}_N \\ \mathbf{1}_N^T & 0 \end{bmatrix} \begin{bmatrix} \alpha_k \\ \lambda_k \end{bmatrix} = \begin{bmatrix} \mathbf{0}_N \\ 1 \end{bmatrix}. \quad (21)$$

式中

$$\mathbf{0}_N = [0, 0, \dots, 0]^T.$$

根据矩阵求逆公式^[3]

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} I_{n1} & -A_{11}^{-1}A_{12} \\ 0 & I_{n2} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix} \begin{bmatrix} I_{n1} & 0 \\ -A_{21}A_{11}^{-1} & I_{n2} \end{bmatrix}, \quad (22)$$

可得

$$\begin{bmatrix} \hat{\alpha}_k \\ \hat{\lambda}_k \end{bmatrix} = \begin{bmatrix} P_k \mathbf{1}_N (\mathbf{1}_N^T P_k \mathbf{1}_N)^{-1} \\ (\mathbf{1}_N^T P_k \mathbf{1}_N)^{-1} \end{bmatrix}. \quad (23)$$

式中

$$P_k = \left(\sum_{k=1}^L E_k E_k^T \right)^{-1}. \quad (24)$$

令 p_{ij} 是 P_k 中的元素, $i = 1, 2, \dots, N; j = 1, 2, \dots, N$, 则可得

$$\alpha_k = \left[\sum_{j=1}^N p_{1j}, \sum_{j=1}^N p_{2j}, \dots, \sum_{j=1}^N p_{Nj} \right]^T / \left[\sum_{i=1}^N \sum_{j=1}^N p_{ij} \right], \quad (25)$$

$$\lambda_k = 1 / \left[\sum_{i=1}^N \sum_{j=1}^N p_{ij} \right]. \quad (26)$$

如果我们把一个矩阵的质量定义为矩阵中所有元素之和,把矩阵的行分质量定义为矩阵中每行元素之和,则内插因子 α_k 的物理意义就是矩阵 P_k 的行质量与矩阵总质量的各个比值构成的向量.所以,可把内插因子视为权重,它表示多模态模型中每个子模型所占的分量.

为了防止数据饱和,也可引进带有遗忘因子的参数辨识方法.

4 仿真研究

考虑下列双线性系统

$$y_k = 0.8y_{k-1} + 0.1y_{k-2}u_{k-1} + u_{k-2} + \varepsilon_k. \quad (27)$$

式中 ε_k 是零均值白噪声,其方差为 $\sigma = 1.0 \times 10^{-4}$.

假设我们将要建立的多模态模型的模态数为 $N = 3$. 取如下数值的三个平衡点

$$(y^*, u^*) = \{(-5, -7.14), (-4, -6.67), (-3, -6)\}.$$

容易算得对应于这三个平衡点的线性化模型分别为

$$\Delta y_k^1 = 0.8\Delta y_{k-1}^1 - 0.5\Delta y_{k-2}^1 - 0.71\Delta u_{k-1}^1 + \Delta u_{k-2}^1, \quad (28)$$

$$\Delta y_k^2 = 0.8\Delta y_{k-1}^2 - 0.4\Delta y_{k-2}^2 - 0.67\Delta u_{k-1}^2 + \Delta u_{k-2}^2, \quad (29)$$

$$\Delta y_k^3 = 0.8\Delta y_{k-1}^3 - 0.3\Delta y_{k-2}^3 - 0.6\Delta u_{k-1}^3 + \Delta u_{k-2}^3. \quad (30)$$

假设双线性系统的输入 u_k 按照图 2 所示形式变化,则输出将如图 3 形式变化.

按照上节所给出的参数辨识方法,我们得到了图 4~6 的结果.在仿真中遗忘因子取为 0.9.

在图 4~6 中,也给出了根据输入 u 按分段线性方式计算得到的内插因子 $\bar{\alpha}_k$ 值.从这些图上看,内插因子值是时间 t 的函数,亦即采样次数 k 的函数.然而,从本质上看,它们的变化起因于 u_k 随时间的变化,同时,不同的 u_k 对应于不同的工作点,模型参数 α_i 应可认为是相应设定值 r 的函数.

图 7 给出了双线性系统(式 27)的输出 y_k 和多模态模型的输出 $\hat{y}_k$ 的差值(可称为模型偏差)的变化曲线.

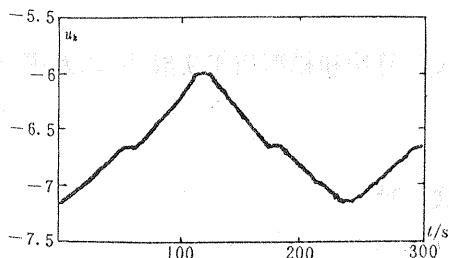


图 2 输入变化曲线

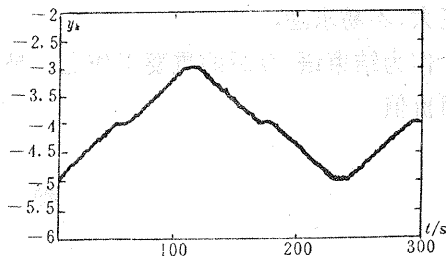
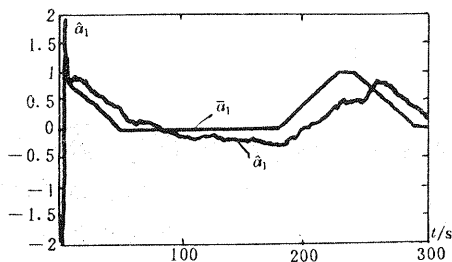
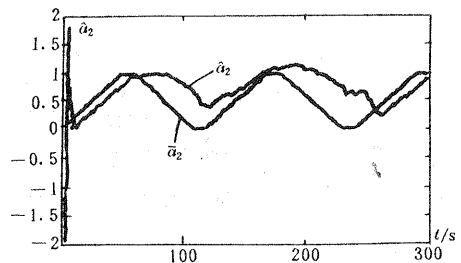
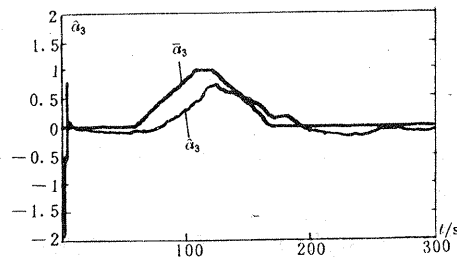
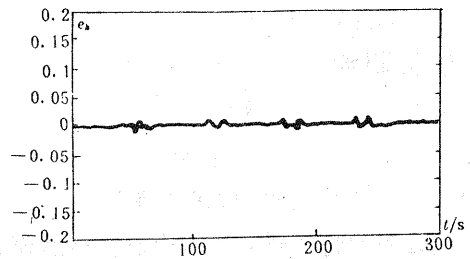


图 3 输入变化曲线

图 4 估计值 $\hat{\alpha}_1$ 图 5 估计值 $\hat{\alpha}_2$ 图 6 估计值 $\hat{\alpha}_3$ 图 7 模型偏差 e_2

5 讨 论

1) 从图 7 的结果看,模型偏差在输入 u 的整个变化过程中和整个参数辨识过程中近似为零,这说明多模态模型是极其有效的,同时,参数辨识算法也是有效的。

2) 从图 4~6 的结果看,由参数辨识方法得到的内插因子值与按分段线性方式计算得到的内插因子变化曲线的走势有一定的相似性。经过比较发现,前者的模型偏差更小。在精度要求不高的场合,采用分段线性方法来计算内插因子,这样可避免辨识过程。

3) 通过大量的仿真,结果表明要使参数辨识获得满意的结果,必须处理好以下几个因素:

i) 遗忘因子的选取,遗忘因子数值越小,待估计参数收敛越快,但对外界扰动也越敏感;

ii) 输入 u 的变化速率,此速率不能过快,否则会对辨识算法造成很大干扰;

iii) 工作点的取法,须注意使各个线性化模型(例如式(28)~(30))间的变化不能太大,否则易使式(12)的误差向量 E_k 的各个分量之间差别甚大,因而使式(24)中的矩阵刚度变大,不易求逆。

作为结束语,今后的重要工作之一是将本文的算法和模型用于实践中,以发挥方法的应用价值。

参 考 文 献

- [1] 周超俊,蒋慰孙,藤井省三. 非线性系统的多模态 ARMAX 模型及其应用. 华东化工学院学报, 1992, 18(增刊): 124-129
- [2] 周超俊,蒋慰孙,藤井省三. 非线性系统的多模态状态模型及其应用. 河海大学学报, 1993, 20(管理专辑): 29-

[3] 倪国熙. 常用的矩阵理论和方法. 上海: 上海科技出版社, 1984, 61

The Multiple Mode ARMAX Model for Nonlinear System and Its Parameter Identification

ZHOU Chaojun and JIANG Weisun

(Research Institute of Automatic Control, East China University of Science and Technology • Shanghai, 200237, PRC)

S. Fujü

(Department of Information Engineering, Nagoya University • Nagoya, Japan)

Abstract: A multiple mode ARMAX model with its parameters interpolated is developed by using linearized models at several operating points of the nonlinear system. Such a model is more accurate in a larger working extent than the traditional linearized model. The identification problem of a MM ARMAX model is also studied. Simulation results have demonstrated the effectiveness of the MM model and the parameter identification algorithm of MM ARMAX model.

Key words: nonlinear system; parameter identification; multiple mode model

本文作者简介

周超俊 1961 年生, 于 1984 年, 1987 年与 1992 年在华东化工学院分别获得生产过程自动化专业学士、工业自动化专业硕士与博士学位. 其间作为中日联合培养的博士生, 在日本名古屋大学得到藤井省三教授的指导. 主要研究领域为多变量控制、系统辨识、故障诊断与智能控制等.

蒋慰孙 1926 年生, 于 1947 年毕业于上海交通大学. 1978 年起在华东化工学院任教授, 1981 年起为博士生导师. 现担任中国自动化学会常务理事、国务院学位委员会学科评议组成员等. 主要研究领域为过程模型化、控制及优化, 近年来并在容错控制、智能控制、柔性过程系统等领域开辟新的研究.