

文章编号: 1001-0920(2011)05-0761-07

基于自适应遗传算法的多项式模型结构与参数的一体化辨识

翟旭升, 王海涛, 谢寿生, 苗卓广, 吴勇

(空军工程大学 工程学院, 西安 710038)

摘要: 基于自适应遗传算法, 提出一种多项式模型结构与参数的一体化辨识方法. 针对组合非线性系统, 首先将选定的候选项原始序列与输出序列进行相关度评估, 根据其大小排列进行遗传算法染色体结构的自适应编码; 在迭代辨识充分后, 再次计算候选项贡献序列与由该项造成的模型损失序列间的相关度, 剔除相关度较小的项, 调整模型结构; 如此循环迭代, 在完成参数辨识的同时最终确认模型结构. 仿真实例验证了算法的有效性.

关键词: 非线性系统辨识; 自适应遗传算法; 相关系数; 结构辨识; 多项式模型; 智能计算

中图分类号: TP13

文献标识码: A

Structure validation and parameter estimation for polynomial nonlinear system using adaptive genetic algorithm

ZHAI Xu-sheng, WANG Hai-tao, XIE Shou-sheng, MIAO Zhuo-guang, WU Yong

(Engineering College, The Air Force Engineering University, Xi'an 710038, China. Correspondent: ZHAI Xu-sheng, E-mail: zhai120120@163.com)

Abstract: A method of structure validation and parameter estimation based on adaptive genetic algorithm is proposed. For an assembled polynomial nonlinear system, the correlation coefficient between the selected candidate item's initial sequence and its output sequence is calculated, and the chromosome's structure of genetic algorithm is coded adaptively according to the correlation coefficient array. After that, the correlation coefficient between the candidate item's contribution sequence and the lost sequence that caused by this item is calculated, the items with smaller correlation coefficients are eliminated, and the model's structure is regulated. Finally the model's structure and parameters can be obtained by repeating the above steps. Several simulation results show the effectiveness of the method.

Key words: nonlinear system identification; adaptive genetic algorithm; correlation coefficient; structure validation; polynomial model; intelligent computation

1 引言

随着研究对象复杂度的提高和人们对控制品质需求的不断增长, 反映客观世界本质的非线性系统建模问题逐渐成为学术界研究的热点之一. 如早期涉及的双线性模型, Hammerstein 模型, Wiener 模型, 非线性时间序列模型, 输入输出仿射模型等^[1-3]. 近年来, 具有普适意义上的非线性模型和智能方法也得到了长足发展, 如 Volterra 级数模型, GFRF 频域映射模型, NARMAX 模型, 模糊逻辑模型^[4], 神经网络模型, 支持向量学习机模型^[5], 遗传规划以及遗传基因表达式模型等^[6-7]. 对上述非线性模型进行深入分析可知, 有一大类模型可以描述为线参数多项式组合非线性模型, 只是模型的多项式表达形式不同. 针对线参

数多项式组合模型的结构辨识与参数辨识方法国内外均进行了研究, 如文献 [8-11] 采用 Gram-Schmidt 正交化方法和 Cholesky 分解方法进行模型的简化工作; [12] 采用双对角正交化方法对 NARMAX 模型进行结构与参数辨识研究; [13] 提出子集优化方法进行多项式模型结构的筛选; [14] 在对输入矩阵进行 Householder 变换过程中, 提出利用输入向量对输出向量的影响因子大小进行结构选择及辨识; [15] 提出递推最小二乘循环迭代方法, 逐步删除那些结构项系数绝对值较小的项, 最终实现结构与参数的辨识.

随着智能优化方法(遗传算法、粒子群算法、差分进化算法)的出现, 采用智能优化方法进行参数辨识已成为研究热点之一. 如文献 [16] 提出一种遗传算

收稿日期: 2010-03-09; 修回日期: 2010-07-19.

作者简介: 翟旭升(1984—), 男, 博士生, 从事航空推进系统综合控制的研究; 谢寿生(1959—), 男, 教授, 博士生导师, 从事航空推进系统综合控制等研究.

法来辨识稀疏 Volterra 系统, 以降低候选项数目; 文献 [17-18] 采用遗传算法及其改进方法进行参数辨识.

本文基于染色体结构自适应遗传算法提出一种结构辨识策略, 对多项式组合非线性模型结构与参数的一体化进行辨识. 即将遗传算法编码自适应于线参数多项式模型的结构, 并依据观测过程中候选项的贡献序列与由该项造成的系统损失序列间的相关度大小确定候选项的重要程度, 在确定模型结构的同时辨识模型参数, 实现结构与参数的一体化辨识.

2 线参数多项式模型描述

普适性的非线性系统有一大类可采用线参数多项式组合模型来统一描述, 其通用表达式为

$$f(x) = \theta_0 + \theta_1 \phi_1(x) + \theta_2 \phi_2(x) + \cdots + \theta_n \phi_n(x). \quad (1)$$

其中: $\theta_i (i = 0, 1, \dots, n)$ 为模型系数, 该系数未知但有界; $\phi_i(x) (i = 0, 1, \dots, n)$ 为模型多项式, 其个数为有限个, 其形式决定模型类型. 如果 $\phi_i(x) \in \{u(j), u(j) * u(k), u(j) * u(k) * u(l), \dots\}$, $j, k, l = 1, 2, \dots, n$, 其中 $u(j)$ 为系统输入项, 则为 Volterra 级数模型; 如果 $\phi_i(x) \in \{u(j), y(j), u(j) * u(k), u(j) * y(k), y(j) * y(k), \dots\}$, $j, k = 1, 2, \dots, n$, 其中 $u(j)$ 和 $y(j)$ 分别为系统输入、输出项, 则为 NARMAX 模型; 如果 $\phi_i(x) \in \{K(x, x_1), K(x, x_2), \dots, K(x, x_l), b\}$, 其中 x_1, x_2, \dots, x_l 表示 l 个训练样本, $K(\cdot)$ 表示样本输入的非线性映射, 则上述模型可以看成基于核函数映射的非线性模型, 当系数具有一定的特殊依赖关系时, 该模型表示为支持向量机模型; 如果 $\phi_i(x) \in \{(x, \sin x, \sqrt{x}, \ln x, b, \dots), [+ , * , / , \dots]\}$, 其中 x 为输入向量, (\cdot) 表达输入向量的非线性描述, $[\cdot]$ 表示非线性运算符, 则模型为基因表达式模型, 均可称为组合非线性多项式模型. 文献 [15] 中, 将非线性模型分为几种特殊形式, 如输入非线性、输出非线性以及混合形式. 实际物理过程中有一大类非线性模型可描述为输入非线性形式, 本文便以输入非线性形式为研究对象. 在辨识过程中, 输入输出数据的处理采用该文中的拓扑同胚映射方法, 将输入输出数据映射于 $[0, 1]$ 区间, 建立非线性模型后, 再通过逆变换还原为原系统.

3 结构辨识策略

对于系统 (1), $y(k) = \sum_{i=1}^N \theta_i * h_i(k)$. 其中: $h_i(k)$ 表示 k 时刻第 i 候选项, θ_i 为贡献权重. 辨识过程中, 设 $\hat{\theta}_i(m)$ 为第 m 步辨识值, 则第 i 项贡献为 $\text{out}_i(k) = \hat{\theta}_i * h_i(k)$, 由该项造成的损失 $\text{loss}_i(k) = y(k) - \sum_{j=1, j \neq i}^N \hat{\theta}_j * h_j(k)$. 因此, 在连续 L 帧采样观测过程中,

第 i 候选项存在两个序列: 贡献序列为 $\text{out}_i(k)$, $k = 1, 2, \dots, L$; 系统损失序列为 $\text{loss}_i(k)$, $k = 1, 2, \dots, L$. 它们之间的相关系数为 $\rho(h_i) = \text{Cov}(\text{out}_i, \text{loss}_i) / \sqrt{\text{Var}(\text{out}_i) \text{Var}(\text{loss}_i)}$.

当系统不受干扰时, 在模型逼近实际系统过程中, 如果第 i 项是系统结构项, 则 $\rho(h_i)$ 随着迭代过程的收敛趋近于 1; 如果第 i 项不是系统结构项, 则 $\rho(h_i)$ 随着迭代过程的收敛趋近于 0. 当系统输入输出受到独立噪声干扰时, $\rho(h_i)$ 小于 1, 在模型逼近实际系统的过程中, 如果第 i 项是系统结构项, 则 $\rho(h_i)$ 随着迭代过程的收敛接近于一个稳定值, 该稳定值受噪声强度影响; 如果第 i 项不是结构项, 则 $\rho(h_i)$ 随着迭代过程的收敛趋近于 0, 且小于系统结构项的相关系数.

引理 1 假设 $\xi(k) (k = 1, 2, \dots, L)$ 为一个随机序列, $E\xi = \xi_0$, $\sigma(\xi) = \sigma_\xi$, 则 $\rho(\xi, \xi) = 1$, 即自相关系数为 1^[19].

定理 1 假设 $\xi(k) (k = 1, 2, \dots, L)$ 为一个随机序列, $E\xi = \xi_0$, $\sigma(\xi) = \sigma_\xi$; $\text{Noi}(k) (k = 1, 2, \dots, L)$ 为方差为 $\sigma(\text{Noi})$ 的独立噪声序列, 则 $\rho(\xi, \xi + \text{Noi}) < \rho(\xi, \xi) = 1$.

证明

$$\begin{aligned} \rho(\xi, \xi + \text{Noi}) &= \frac{E[(\xi - E\xi)(\xi + \text{Noi} - E(\xi + \text{Noi}))]}{\sigma(\xi)\sigma(\xi + \text{Noi})} = \\ &= \frac{E[(\xi - E\xi)(\xi + \text{Noi} - E\xi)]}{\sqrt{E\xi^2 - (E\xi)^2} \sqrt{E(\xi + \text{Noi})^2 - (E(\xi + \text{Noi}))^2}} = \\ &= \frac{E(\xi^2) - \xi_0^2}{\sqrt{E(\xi^2) - \xi_0^2} \sqrt{E(\xi^2) + (\sigma(\text{Noi}))^2 - \xi_0^2}} < 1. \end{aligned}$$

定理得证. \square

定理 2 假设 $\xi(k) (k = 1, 2, \dots, L)$ 为一个随机序列, $E\xi = \xi_0$, $\sigma(\xi) = \sigma_\xi$; $\text{Noi}_1(k) (k = 1, 2, \dots, L)$ 为方差为 $\sigma(\text{Noi}_1)$ 的独立噪声序列, $\text{Noi}_2(k) (k = 1, 2, \dots, L)$ 为方差为 $\sigma(\text{Noi}_2)$ 的独立噪声序列, 则 $\rho(\xi + \text{Noi}_1, \xi + \text{Noi}_2) < \rho(\xi + \text{Noi}_1, \xi) < 1$.

证明

$$\begin{aligned} \rho(\xi + \text{Noi}_1, \xi + \text{Noi}_2) &= \frac{E[(\xi + \text{Noi}_1 - E(\xi + \text{Noi}_1))(\xi + \text{Noi}_2 - E(\xi + \text{Noi}_2))]}{\sigma(\xi + \text{Noi}_1)\sigma(\xi + \text{Noi}_2)} = \\ &= \frac{E(\xi^2) - \xi_0^2}{\sqrt{E(\xi^2) + (\sigma(\text{Noi}_1))^2 - \xi_0^2} \sqrt{E(\xi^2) + (\sigma(\text{Noi}_2))^2 - \xi_0^2}} < \\ &= \frac{E(\xi^2) - \xi_0^2}{\sqrt{E(\xi^2) - \xi_0^2} \sqrt{E(\xi^2) + (\sigma(\text{Noi}_2))^2 - \xi_0^2}} = \\ \rho(\xi + \text{Noi}_1, \xi) &< 1. \end{aligned}$$

定理得证. \square

推论 1 假设 $\xi(k) (k = 1, 2, \dots, L)$ 是一个随机

序列, $E\xi = \xi_0, \sigma(\xi) = \sigma_\xi$; $\text{Noi}_1(k)$ 是方差为 $\sigma(\text{Noi}_1)$ 的随机序列, $\text{Noi}_2(k)$ 是方差为 $\sigma(\text{Noi}_2)$ 的随机序列, $\text{Noi}'_1(k)$ 是方差为 $\sigma(\text{Noi}'_1)$ 的随机序列, $\text{Noi}'_2(k)$ 是方差为 $\sigma(\text{Noi}'_2)$ 的随机序列, Noi_2 与 Noi_1 不相关, Noi'_2 与 Noi'_1 不相关, 且满足 $\sigma(\text{Noi}'_2) < \sigma(\text{Noi}_2), \sigma(\text{Noi}'_1) < \sigma(\text{Noi}_1)$. 则有 $\rho(\xi + \text{Noi}_1, \xi + \text{Noi}_2) < \rho(\xi + \text{Noi}'_1, \xi + \text{Noi}'_2)$.

引理2 在模型输入输出受白噪声干扰时, 迭代过程是一致收敛的, 即在迭代步骤 $m \rightarrow \infty$ 时, 其模型辨识参数 $\hat{\theta}_m \xrightarrow[m \rightarrow \infty]{W.P.1} \theta_0$, 或称模型参数以概率1收敛于最优值, 即真值^[20].

定理3 结构辨识策略: 在迭代逐步收敛的过程中, 模型结构项的贡献序列与由该结构项造成的系统损失序列间的相关度概率性有所提高, 且逐步收敛于一个稳定值.

证明 定义一组离散序列 $y(k), y_{ob}(k), \text{Noi}_{ob}(k)$ 分别为系统真实值、测量值、测量噪声, $k = 1, 2, \dots, L$. 假设测量噪声是方差大于零的随机序列; $h_i(k) (k = 1, 2, \dots, L, i = 1, 2, \dots, N)$ 是第 i 候选项的 k 时刻采样, N 是候选项个数; $\theta_i (i = 1, 2, \dots, N)$ 是候选项权重.

由上述定义可知, $y(k) = \sum_{i=1}^N \theta_i * h_i(k), y_{ob}(k) = y(k) + \text{Noi}(k), k = 1, 2, \dots, L$. 设第 m 代辨识的模型参数为 $\hat{\theta}_i(m), i = 1, 2, \dots, N$; 第 $m+1$ 代辨识的模型参数为 $\hat{\theta}_i(m+1), i = 1, 2, \dots, N$. $\hat{y}_m(k)$ 表示由第 m 代辨识模型的预测值序列, 满足 $\hat{y}_m(k) = \sum_{i=1}^N \hat{\theta}_i(m) * h_i(k)$; 同理 $\hat{y}_{m+1}(k)$ 表示由第 $m+1$ 代辨识模型的预测值序列, 满足 $\hat{y}_{m+1}(k) = \sum_{i=1}^N \hat{\theta}_i(m+1) * h_i(k)$. 则 $y(k)$ 可表示如下:

$$\begin{aligned} y(k) &= \sum_{i=1}^N \theta_i * h_i(k) = \\ &= \sum_{i=1}^N (\hat{\theta}_i(m) + \text{err}_i(m)) * h_i(k) = \\ &= \hat{y}(k, m) + \sum_{i=1}^N \text{err}_i(m) * h_i(k) = \\ &= \sum_{i=1}^N (\hat{\theta}_i(m+1) + \text{err}_i(m+1)) * h_i(k) = \\ &= \hat{y}(k, m+1) + \sum_{i=1}^N \text{err}_i(m+1) * h_i(k). \end{aligned}$$

其中: $\hat{\theta}_i(m)$ 表示参数 θ_i 的第 m 代估值, $\text{err}_i(m) = \theta_i - \hat{\theta}_i(m)$ 为估值偏差; 同理 $\hat{\theta}_i(m+1)$ 表示参数 θ_i 的第 $m+1$ 代估值, $\text{err}_i(m+1) = \theta_i - \hat{\theta}_i(m+1)$ 为估

值偏差. 随着迭代过程的进行, $\hat{\theta}_m \xrightarrow[m \rightarrow \infty]{W.P.1} \theta_0$, 因 $\text{abs}(\hat{\theta}_i(m) - \theta_i) > \text{abs}(\hat{\theta}_i(m+1) - \theta_i)$ 在概率意义上成立, 所以 $\text{err}_i(m+1) * h_i(k) < \text{err}_i(m) * h_i(k)$.

由定义可知, 模型共有 N 项, 观测长度为 L ; 第 i 项的贡献序列为 $\theta_i * h_i(k), k = 1, 2, \dots, L$; 迭代过程中, 第 m 代模型第 i 项的贡献序列为 $\hat{\theta}_i(m) * h_i(k)$; 第 $m+1$ 代模型第 i 项的贡献序列为 $\hat{\theta}_i(m+1) * h_i(k)$. 由于在辨识过程中不知道第 i 项实际贡献序列, 可采用模型在缺少第 i 项贡献时的系统损失序列 $\text{loss}_i(k)$ 来代替. 系统损失序列如下所示:

$$\begin{aligned} \text{loss}_i(k) &= y_{ob}(k) - (\hat{y}(k) - \hat{\theta}_i * h_i(k)) = \\ &= \hat{\theta}_i * h_i(k) + \sum_{i=1}^{LS} \text{err}_i * h_i(k) + \text{Noi}_{ob}(k). \end{aligned} \quad (2)$$

此处, 将第 m 代模型第 i 项的系统损失序列定义为 $\text{loss}_i(k, m)$, 满足

$$\begin{aligned} \text{loss}_i(k, m) &= \\ &= \hat{\theta}_i(m) * h_i(k) + \sum_{i=1}^N \text{err}_i(m) * h_i(k) + \text{Noi}_{ob}(k). \end{aligned} \quad (3)$$

同理, 第 $m+1$ 代模型第 i 项的系统损失序列定义为 $\text{loss}_i(k, m+1)$, 满足

$$\begin{aligned} \text{loss}_i(k, m+1) &= \\ &= \hat{\theta}_i(m+1) * h_i(k) + \\ &= \sum_{i=1}^N \text{err}_i(m+1) * h_i(k) + \text{Noi}_{ob}(k). \end{aligned} \quad (4)$$

其中: $\text{Noi}_{ob}(k)$ 为零均值的观测白噪声序列; $\sum_{i=1}^N \text{err}_i(m) * h_i(k), \sum_{i=1}^N \text{err}_i(m+1) * h_i(k)$, 分别为第 $m, m+1$ 代辨识模型输出值与系统真实值之间的偏差序列, $k = 1, 2, \dots, L$, 记为 $\sigma(m), \sigma(m+1)$; 由 $\text{abs}(\hat{\theta}_i(m) - \theta_i) > \text{abs}(\hat{\theta}_i(m+1) - \theta_i)$ 在统计意义上成立, 所以 $\sigma(m) > \sigma(m+1)$ 也在统计意义上成立.

第 m 代第 i 模型项贡献序列与系统损失序列间的相关系数为 $\rho(\text{out}_i(m), \text{loss}_i(m))$, 其中

$$\begin{aligned} \text{loss}_i(m) &= \\ &= \hat{\theta}_i(m) * h_i(k) + \sum_{i=1}^N \text{err}_i(m) * h_i(k) + \text{Noi}_{ob}(k), \\ \text{out}_i(m) &= \hat{\theta}_i(m) * h_i(k), k = 1, 2, \dots, L. \end{aligned}$$

由以上的推理可知

$$\begin{aligned} \rho(\text{out}_i(m), \text{loss}_i(m)) &= \\ &= \rho(\hat{\theta}_i(m) * h_i(k), \hat{\theta}_i(m) * h_i(k) + \\ &= \sum_{i=1}^N \text{err}_i(m) * h_i(k) + \text{Noi}_{ob}(k)) = \end{aligned}$$

$$\rho\left(\theta_i * h_i(k) - \text{err}_i(m) * h_i(k), \theta_i * h_i(k) - \text{err}_i(m) * h_i(k) + \sum_{i=1}^N \text{err}_i(m) * h_i(k) + \text{Noi}_{ob}(k)\right).$$

由于 $\text{err}_i(m) * h_i(k)$ 与 $\text{Noi}(k)$ ($k = 1, 2, \dots, L$) 是不相关的, 由定理 2 可知

$$\rho(\text{out}_i(m), \text{loss}_i(m)) \leq \rho(\theta_i * h_i(k), \theta_i * h_i(k) + \text{Noi}_{ob}(k)). \quad (5)$$

同理, 第 $m + 1$ 代辨识模型第 i 项的贡献序列与系统损失序列的相关系数为

$$\begin{aligned} & \rho(\text{out}_i(m+1), \text{loss}_i(m+1)) = \\ & \rho\left(\hat{\theta}_i(m+1) * h_i(k), \hat{\theta}_i(m+1) * h_i(k) + \sum_{i=1}^N \text{err}_i(m+1) * h_i(k) + \text{Noi}_{ob}(k)\right) = \\ & \rho\left(\theta_i * h_i(k) - \text{err}_i(m+1) * h_i(k), \theta_i * h_i(k) - \text{err}_i(m+1) * h_i(k) + \sum_{i=1}^N \text{err}_i(m+1) * h_i(k) + \text{Noi}_{ob}(k)\right). \end{aligned}$$

因 $\text{abs}(\text{err}_i(m)) > \text{abs}(\text{err}_i(m+1))$ 在统计意义上是成立的, 所以 $\text{err}_i(m+1) * h_i(k)$ ($k = 1, 2, \dots, L$) 随机序列的标准差要小于 $\text{err}_i(m) * h_i(k)$ ($k = 1, 2, \dots, L$) 随机序列的标准差. 由定理 2 可知

$$\rho(\text{out}_i(m+1), \text{loss}_i(m+1)) > \rho(\text{out}_i(m), \text{loss}_i(m)).$$

随着迭代过程的收敛, $\hat{\theta}_{i,m} \xrightarrow{W.P.1} \theta_i$, $\text{err}_{i,m} \xrightarrow{W.P.1} 0$, 相关系数趋于稳定, 该稳定值为 $\rho_i(\infty)$ 且 $\rho_i(\infty) = \rho(\theta_i * h_i(k), \theta_i * h_i(k) + \text{Noi}_{ob}(k))$. \square

另外由相关系数表达式可知, 观测噪声方差越大, 每一候选项贡献序列与损失序列间的相关系数越小, 非结构项的贡献序列与损失序列间的相关系数更具随机性. 当观测噪声强度一定时, 模型结构项其贡献序列与对应系统损失序列间的相关系数较大, 且显著大于那些非结构项的相关系数; 贡献越大, 相关性越强.

4 自适应遗传算法流程及设计

4.1 算法流程

Step 1: 将输入、输出进行相应的拓扑同胚映射, 映射区间为 $[0, 1]$; 确定所有的结构候选项, 定义候选项输入序列 $h_i(k)$, 输出序列 $y(k)$. 对于任何的有限区间 $[a, b]$, 总存在一个拓扑同胚映射, 即将区间 $[a, b]$ 一一对应地映射到 $[0, 1]$ 区间, 具体实现方法及证明过程见文献 [15].

Step 2: 根据 $\rho(h_i) = \text{Cov}(h_i, y_i) / \sqrt{\text{Var}(h_i)\text{Var}(y_i)}$, 计算所有候选项原始输入序列与输出序列间的相关

度, 并实施逆序排列.

Step 3: 根据逆序的排列, 将所有结构项按照相关度大小分成一个或者几个候选块进行结构自适应编码.

Step 4: 设置基本遗传算法参数, 如种群规模 n , 变异概率 p_m , 交叉概率 p_v , 迭代停止条件, 参数范围初始化, 适应度函数 f 的计算等.

Step 5: 开始循环迭代, 进行相应交叉、变异操作; 当达到停止条件, 转 **Step 6**.

Step 6: 判断. 如果适应度函数 f 大于界限值, 则停止总体迭代, 输出辨识结构及对应参数, 退出循环; 否则, 转 **Step 7**.

Step 7: 结构确认. 由 $\rho(h_i) = \text{Cov}(\text{out}_i, \text{loss}_i) / \sqrt{\text{Var}(\text{out}_i)\text{Var}(\text{loss}_i)}$ 计算每一项的相关度. 如果相关度大于界限阈值 (本文选定为 0.25), 则保留该结构项; 否则, 删除该结构项. 确认完毕, 转 **Step 8**.

Step 8: 将优选结构项与原先的未筛选结构项重新排列, 进行染色体长度的自适应调整和编码, 转 **Step 5**.

需要指出的是, 如果候选项较少, 便可一次加入, 该过程即为一个逐步剔除非结构项的结构辨识特例. 同时停止条件与筛选策略的设定, 应确保在将所有项进行充分激励的同时, 避免将真实模型项误删除.

4.2 适应度函数的设计

遗传算法适应度函数的设置原则为: 1) 在与真实系统比较时, 应能反映辨识模型与被辨识系统间的接近程度; 2) 在结构确认时, 应能阻止那些非结构项的干扰以及测量噪声的干扰, 即能够避免误判非模型项. 以下分两个方面进行讨论:

1) 不受噪声干扰

由于没有噪声, 只需将模型输出与系统输出在观测过程之间均方差最小化即可, 均方差表达式设置如下:

$$J = \frac{1}{L} * \sum_{i=1}^L (y_i - \hat{y}_i)^2. \quad (6)$$

其中: L 为采样观测长度, y_i 和 \hat{y}_i 分别为系统值与模型值. 为了更有效地判断和转化, 对式 (6) 进行适当转换, 采用指数表达式实现这一目标, 定义适应度函数表达式如下:

$$F = \exp(-J * c), \quad (7)$$

其中 $c(c \geq 1)$ 是适应度影响因子.

2) 受高斯噪声干扰

如果系统被白色噪声 $\mu(k)$ 所污染, 且噪声与输入向量 $\mu(k)$ 是非相关的, 则

$$y_t(k) = y(k) + \mu(k), \quad (8)$$

其中 $y(k)$ 是无污染的系统输出. 假定噪声是一个零均值序列, 且其方差 σ_μ^2 已知. 定义误差为 $e(k) = y(k) - \hat{y}(k)$, 则误差最大似然函数为

$$L = \prod_{i=1}^L p(e(i)), p = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp\left(\frac{-e^2(i)}{2\sigma_\mu^2}\right). \quad (9)$$

因为最大似然函数是随机的, 所以对似然函数取对数可以很好地进行估计. 因此, 最大化的目标函数可如下选择:

$$F_1 = \sum_{i=1}^L \log p(e(i)) = \sum_{i=1}^L \log \frac{1}{\sqrt{2\pi\sigma_\mu^2}} - \prod_{i=1}^L \left(\frac{e^2(i)}{2\sigma_\mu^2}\right) = C - \frac{1}{2\sigma_\mu^2} \sum_{i=1}^L e^2(i) = C - \frac{L}{2} \frac{\sigma_e^2}{\sigma_\mu^2}.$$

其中: C 为一个常量, $\sigma_e^2 = \frac{1}{L} \sum_{i=1}^L e^2(i)$ 为辨识模型输出误差方差. 在进化早期, σ_e^2 较大; 随着进化过程的推进, 模型接近真实系统, 此过程中 σ_e^2 将减小并且越来越接近于 σ_μ^2 . 很明显, 因为噪声 σ_e^2 大于零, 而当 σ_e^2 与 σ_μ^2 接近相等时, 可认为辨识模型与实际系统最接近. 为此, 定义新的适应度函数如下:

$$F_u = \exp(-|\sigma_e^2 - \sigma_\mu^2|). \quad (10)$$

可见, 在模型充分接近系统的过程中, 适应度函数同样不断趋近于 1.

5 仿真实例

为验证算法有效性, 本节选用两个典型非线性系统进行仿真, 平台为主频 2.4 GHz, 内存为 256 M 的 PC 机, 选择软件为 Matlab7.0.

例 1 假定模型结构项表达式为

$$y(k) = 0.8u(k-1) + 1.4u(k-2) + 0.9u(k-3) + 0.5u(k-4) - 2.0u(k-1)u(k-1) + u(k-1)u(k-2) - 1.3u(k-1)u(k-3) + u(k-1)u(k-4) - 1.5u(k-1)u(k-5) + u(k-2)u(k-2) + 1.8u(k-2)u(k-4) + \text{Noi}(k),$$

其中 $u(k)$, $y(k)$, $\text{Noi}(k)$ 分别为 k 时刻的输入、输出和系统干扰噪声. 假定输入 $u(k)$ 为 $[-1, 1]$ 之间的白噪声信号, $\text{Noi}(k)$ 为相对于输入信号强度 1/10 的干扰. 由表达式可知, 系统为 2 阶时延为 5 的稀疏 Volterra 级数系统, 且结构项为 11 项. 在仿真时, 设置系统阶次为 2, 时延为 5, 则结构项共有 20 项.

本实例进行满集结构辨识仿真, 将该过程描述为一个逐步删减过程. 根据上述算法进行仿真, 仿真过

程适应度变化示意图和模型结构项的筛选过程分别如图 1 和图 2 所示, 表 1 为对应辨识结果.

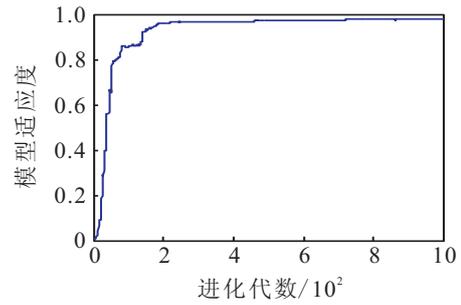


图 1 例 1 的适应度变化趋势

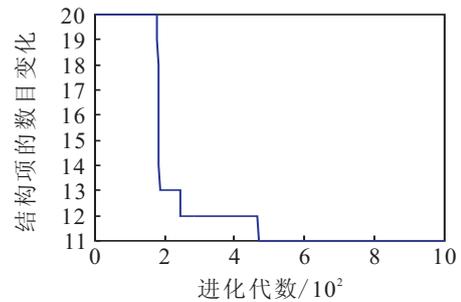


图 2 例 1 的结构项个数变化趋势

表 1 辨识参数对比表

结构项	实际值	辨识值
θ_1	0.8	0.8130
θ_2	1.4	1.3795
θ_3	0.9	0.9135
θ_4	0.5	1.5131
θ_6	-2.0	-1.8748
θ_7	1.0	0.9915
θ_8	-1.3	-1.3148
θ_9	1.0	1.0311
θ_{10}	-1.5	-1.4731
θ_{11}	1.0	0.9337
θ_{13}	1.8	1.7858

例 2 当结构项较多时, 按照初始相关度大小排列分多块进行迭代辨识, 即在进化当前代的(最优个体)整体适应度达到设定值后, 启动删除线程, 删除那些非典型项; 随后在保留上一代优胜结构项的基础上, 将剩余结构项按照相应次序补充, 再次进行编码、迭代进化. 如此循环, 直到系统最优结构能够满足实际要求, 跳出循环.

基于一个简单实例, 表达式如下:

$$y(k) = 1.8u(k) + 1.4u(k-1) + 0.9u(k-2) + u(k)u(k) + 1.3u(k)u(k-2) + u(k)u(k-3) + 1.0u(k-1)u(k-1) + 1.2u(k-1)u(k-3) + 2.3u(k-2) * u(k-3) - 2.3u(k-3)u(k-4) + \text{Noi}(k).$$

该系统是一个时延为 5, 阶次为 2 的系统, 共有

20个结构项. 其中 $u(k)$, $y(k)$, $Noi(k)$ 分别为 k 时刻的输入、输出和系统干扰噪声. 假定输入 $u(k)$ 为 $[-1, 1]$ 之间的白噪声信号, $Noi(k)$ 为相对于输入信号强度1/10的干扰噪声.

此时按照原始信号与输出的相关度大小进行排序, 第1次进化选择8个结构项, 然后依次选入6个, 这样迭代3次即可将系统所有结构项进行筛选. 系统辨识仿真过程如图3和图4所示, 仿真结果如表2所示.

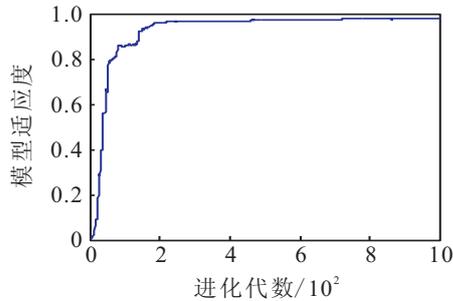


图3 例2的适应度变化趋势

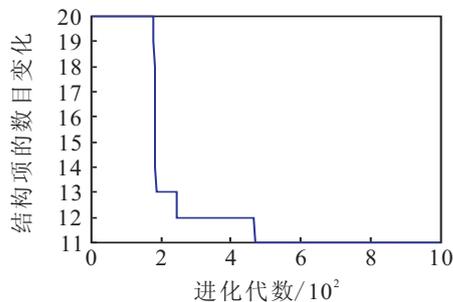


图4 例2的结构项个数变化趋势

表2 例2的结构项参数辨识结果

结构项	实际值	辨识值
θ_1	1.8	1.8124
θ_2	1.4	1.3961
θ_3	0.9	0.8697
θ_6	1.0	0.9624
θ_8	1.3	1.2474
θ_9	1.0	0.9320
θ_{11}	1.0	0.9757
θ_{13}	1.2	1.2101
θ_{16}	2.3	2.2814

6 结论

本文提出了一种基于相关度大小进行候选项确认的结构辨识策略, 结合结构自适应遗传算法实现线参数多项式模型结构与参数的一体化辨识. 对组合非线性系统, 该策略首先将有界输入输出数据拓扑同胚映射于 $[0, 1]$ 区间; 然后将选定的候选项原始序列与输出序列进行相关度评估, 根据其大小排列进行遗传算法的结构自适应编码. 在采用遗传算法迭代辨识参数后, 再次计算候选项贡献序列与由该项造成的模型

损失序列间的相关度, 剔除相关度较小的项. 随后在保留上一代优胜者的基础上, 将剩余结构项按照相应次序补充, 如此循环迭代该过程, 最终实现了线参数多项式模型的结构与参数辨识的一体化. 该方法简单易行, 大量仿真也验证了该方法的有效性.

本文方法在如下几个方面还有待进一步研究: 如何对观测噪声强度或噪声特征未知的系统进行辨识? 在结构辨识时, 相关度阈值与干扰强度有何定性与定量的影响?

参考文献(References)

- [1] Schetzen M. The volterra and wiener theories of nonlinear systems[M]. New York: Willy, 1980.
- [2] 朱全民. 非线性系统辨识[J]. 控制理论与应用, 1994, 11(6): 641-652.
(Zhu Q M. Nonlinear system identification[J]. Control Theory & Applications, 1994, 11(6): 641-652.)
- [3] Leontaritis I J, Billings S A. Input-output parametric models for nonlinear system, Part II: Stochastic nonlinear system[J]. Int J of Control, 1985, 41(2): 1863-1878.
- [4] Takagi T, Sugeno M. Fuzzy identification of systems and its application to modeling and control[J]. IEEE Trans on Systems, Man and Cybernetics, 1985, 15(1): 116-132.
- [5] 相征, 张太镒, 孙建成. 基于最小二乘向量机的非线性系统建模[J]. 系统仿真学报, 2006, 18(9): 2684-2687.
(Xiang Z, Zhang T Y, Sun J C. Modeling of nonlinear systems based on recurrent least squares support vector machines[J]. J of System Simulation, 2006, 18(9): 2684-2687.)
- [6] Katya Rodriguez-Vazquez, Peter J Fleming. Use of genetic programming in the identification of rational model structures[C]. European Conf on Genetic Programming. Scotland, 2000: 181-192.
- [7] Mark P Hinchliffe, Mark J Willis. Dynamic systems modeling using genetic programming[J]. Computers & Chemical Engineering, 2003, 27(12): 1841-1854.
- [8] Billings S A, Korenberg M, Chen S. Identification of nonlinear output-affine systems using an orthogonal least-squares algorithm[J]. Int J of Systems Science, 1988, 19(8): 1559-1568.
- [9] Billings S A, Chen S, Korenberg M J. Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator[J]. Int J of Systems Science, 1989, 49(6): 2157-2189.
- [10] Wei H L, Billings S A, Liu J. Term and variable selection for nonlinear system identification[J]. Int J of Control, 2004, 77(1): 86-110.
- [11] Korenberg M J. A robust orthogonal algorithm for system identification and time-series analysis[J]. Biological

- Cybernetics, 1989, 60(4): 267-276.
- [12] 王晓, 谢剑英, 贾青. 非线性NARMAX模型结构与参数一体化辨识的改进算法[J]. 信息与控制, 2000, 29(2): 102-110.
(Wang X, Xie J Y, Jia Q. New modified intergated algorithm for structure determination and parameter estimation for nonlinear stochastic systems[J]. Information and Control, 2000, 29(2): 102-110.)
- [13] 唐晓泉, 王文正, 方洋旺. 子集优化在非线性系统辨识中的应用[J]. 西安交通大学学报, 1999, 33(3): 19-22.
(Tang X Q, Wang W Z, Fang Y W. Nonlinear system identification by subset optimization[J]. J of Xi'an Jiaotong University, 1999, 33(3): 19-22.)
- [14] 欧文, 韩崇昭, 王文正. Volterra 泛函级数在非线性系统辨识中的应用[J]. 控制与决策, 2002, 17(2): 239-242.
(Ou W, Han C Z, Wang W Z. Application of volterra series in the identification of nonlinear system[J]. Control and Decision, 2002, 17(2): 239-242.)
- [15] 罗秋滨, 李秀英, 韩志刚, 等. 一种非线性系统多项式逼近的建模方法[J]. 系统仿真学报, 2008, 20(4): 829-832.
(Luo Q B, Li X Y, Han Z G, et al. Polynomial approach for nonlinear modeling[J]. J of System Simulation, 2008, 20(4): 829-832.)
- [16] Leechter Yao. Genetic algorithm based identification of nonlinear system by spare volterra filters[J]. IEEE Trans on Signal Process, 1999, 47(12): 3433-3435.
- [17] Hazem M Abbas, Mohamed M Bayoumi. Volterra-system identification using adaptive real coded genetic algorithm[J]. IEEE Trans on System, Man and Cybernetics, Part A: System and Humans, 2006, 36(4): 671-684.
- [18] 任子武, 伞冶. 自适应遗传算法的改进及其在系统辨识中的应用研究[J]. 系统仿真学报, 2006, 18(1): 41-43.
(Ren Z W, San Y. Improved adaptive genetic algorithm and its application research in parameter identification[J]. J of System Simulation, 2006, 18(1): 41-43.)
- [19] 同济大学数学教研室. 概率论[M]. 北京: 高等教育出版社, 1982.
(Mathematical Teaching and Research Section of Tongji University. Probability theory[M]. Beijing: The Higher Education Press, 1982.)
- [20] 黄友锐. 智能优化算法及其应用[M]. 北京: 国防工业出版社, 2008.
(Huang Y R. Intelligent optimization algorithm and its application[M]. Beijing: National Defense Industry Press, 2008.)

(上接第756页)

参考文献(References)

- [1] Pawlak Z. Rough sets[J]. Int J of Information and Computer Science, 1982, 10(11): 341-356.
- [2] Pawlak Z. Rough sets theory and its application to data analysis[J]. Int J of Cybernetics and Systems, 1998, 29(9): 661-688.
- [3] Duntsch I, Gediga G. Simple data filtering in rough set systems[J]. Int J of Approximate Reasoning, 1998, 18(2): 93-106.
- [4] Cho Y, Lee K, Yoo J, et al. Autogeneration of fuzzy rules and membership functions for fuzzy modeling rough set theory[J]. IEE Proc of Control Theory Application, 1998, 145(5): 437-442.
- [5] Ziarko W. Variable precision rough set mode[J]. J of Computer and Systems Science, 1993, 46(1): 39-59.
- [6] 关欣, 衣晓, 孙迎丰, 等. 变精度粗糙集模型及其在辐射源识别中的应用[J]. 清华大学学报, 2007, 47(1): 28-31.
(Guan X, Yi X, Sun Y F, et al. Variable precision rough set model with applications to emitter recognition[J]. J of Tsinghua University, 2007, 47(1): 28-31.)
- [7] 关欣, 衣晓, 何友. 一种新的粗糙集属性约简方法及其应用[J]. 控制与决策, 2009, 24(3): 464-467.
(Guan X, Yi X, He Y. Knowledge reduction and its applications based on rough set[J]. Control and Decision, 2009, 24(3): 464-467.)
- [8] Chen Ting, Luo Jingqing. A fuzzy recognition method of emitter based on variable precision Rough Set model[C]. The 9th Int Conf on Signal Processing Proceedings. Beijing, 2008, 10: 2242-2245.
- [9] 王杰贵, 靳学明, 罗景青. 基于ESM与ELINT信息融合的机载辐射源识别[J]. 电子学报, 2006, 34(3): 424-428.
(Wang J G, Jin X M, Luo J Q. Airborne emitter recognition based on multisensor data fusion of ESM and ELINT[J]. Acta Electronica Sinica, 2006, 34(3): 424-428.)
- [10] 陈功, 张雄伟, 李耀波. 灰关联分析法在坦克声目标识别系统中的应用[J]. 探测与控制学报, 2006, 28(3): 39-42.
(Chen G, Zhang X W, Li Y B. The application of tank acoustic target recognition system based on gray correlation analysis[J]. J of Detection & Control, 2006, 28(3): 39-42.)