文章编号: 1002-0411(2001)03-231-03

基于神经模糊方法的复杂系统建模

李 波 张世英
(天津大学管理学院 300072)

摘 要: 本文提出一种基于模糊模型同径向基函数网络相结合的复杂系统建模方法. 该方法表明具有确定后件结构的 MTS 模糊模型与径向基函数网络之间有一种直接对应关系, 基于这种对应, 我们可把 MTS 模型的前件结构确定和后件结构辨识分开, 利用径向基函数网络的学习特性和其它学习算法相结合来得到模糊模型. 该方法简单且能达到较高精度. 仿真实例说明了所提方法的有效性.

关键词: MTS 模糊模型; 径向基函数网络; 复杂系统; 建模
中图分类号: TP13 文献标识码: B

COMPLEX SYSTEMS MODELING VIA NEURO-FUZZY METHOD

LI Bo ZHANG Shi-ying
(*Management School, Tianjin University, Tianjin* 300072)

Abstract: Complex systems modeling method based on the combination of MTS fuzzy model with the radial basis function is presented in this paper. A direct correspondence is shown between MTS fuzzy model with the determining consequence structures and the radial basis function networks,. The premise structures of MTS model and consequence parameters can be separated in term of this direct correspondence, So we can obtain the fuzzy model using the combination of the radial basis function learning and the square recursive algorithm. The proposed approach has the advantages of simplicity, flexibility and high accuracy. Simulation studies are presented to illustrate the effectiveness of the method.

Keywords: MTS fuzzy model, radial basis function, complex systems, modeling

1 引言(Introduction)

TS 模糊模型是 Takagi 和 Sugeno[1]发展的对研究具有复杂过程、高非线性特性和系统部分未知等方面最具吸引力的方法且已经取得许多研究成果[2,3]. 最近有学者在 TS 模糊模型的基础上, 为减少计算量, 提出一种 TS 模型的变形, 称之为 MTS (Modified TS) 模型[4], 其一般形式为:

$$R^i\text{: IF } X \text{ is } A^i,\ \text{THEN } y^i = f^i(X)$$
$$i = 1, \cdots, m \qquad (1)$$

其中: X 是 n 维输入变量, y^i 是基于规则 R^i 的模型的局部输出, A^i 是具有隶属度函数 $\mu_A(X)$ 的 n 维模糊子集, 显然隶属度函数的参数是前件结构中需辨识的参数, m 是规则个数, $f^i(X)$ 是输入变量的函数, 一般常用的是常数或输入的线性组合形式.

给定某输入 $X = [x_1, x_2, \cdots, x_n]^T$, 总的模型输出由各规则子输出的加权平均来定义:

$$y = \sum_{i=1}^{m} \mu_{A^i}(X) f^i(X) \text{ 且} \sum_{i=1}^{m} \mu_{A^i}(X) = 1 \qquad (2)$$

由此可知, MTS 模型中的每个规则所代表的模糊子集是以一个在 n 维空间中具有一些需调整前件参数的隶属度函数形式表示的. 这样, 在由原 TS 模型中计算各规则的 n 个分量的模糊域隶属度函数问题就可简化为只计算 MTS 模型中各规则的一个隶属度问题, 因此大大减少了计算量.

目前径向基函数网络(the radial basis function, RBF)是最被关注的一种网络模型, 关于它的研究不仅在实际应用上取得许多成果而且在理论上也得到许多有价值的结论[5]. 一般地, 实际中较常用的是 RBF 网络的规范化形式, 称之为 NRBF 网络, 其模型形式为:

收稿日期: 2000-03-30
基金项目: 国家自然科学基金资助项目(No. 69874028)

$$f_m(X)=\frac{\sum_{i=1}^{m}\omega_i\Psi[(X-C_i)^T\Sigma^{-1}(X-C_i)]}{\sum_{i=1}^{m}\Psi[(X-C_i)^T\Sigma^{-1}(X-C_i)]} \quad (3)$$

显然, 它是非线性基函数 $\Psi(\cdot):R^+\to R$ 的线性加权形式, 经研究表明, 非线性函数 $\Psi(\cdot)$ 的选择对 NRBF 网络的性能不很重要, 一般可假设非线性函数 $\Psi(\cdot)$ 是固定的, 如取高斯函数 $\Psi(r)=e^{-r^2}$. 因此, 由公式(3), 可知有三个参数需要估计: 中心点 $C_i\in R^n$, 加权值 ω_i 和尺度矩阵 $\Sigma_{\max}$. 为简化问题, 目前常假设 Σ 是已知的, 如取 $\Sigma=I$ 或 $\Sigma=\sigma(n)^2I$ 且 $\sigma(n)$ 已知, 而中心点 C_i 可直接从样本输入数据集中选取. 在这种假设下, 取得 N 组样本输入 $X(t)$ 和相应的期望输出 $y(t)$, $t=1,\cdots,N$, 加权值 ω_i 由线性最小二乘方法可进行估计. 因此, 辨识 NRBF 网络结构的关键是中心点 C_i 的选择问题.

若我们假设(1)式表示的 MTS 模糊模型的后件结构已知, 即各规则的 y^i 是确定的, 比较(2)式和(3)式, 若取隶属度函数为

$$\mu_{A^i}(X)=\frac{|\Psi[X-C_i)^T\Sigma^{-1}(X-C_i)]|}{\sum_{i=1}^{m}|\Psi[X-C_i)^T\Sigma^{-1}(X-C_i)]|}$$
$$i=1,\cdots,m \quad (4)$$

则 $\sum_{i=1}^{m}\mu_{A^i}(X)=1$

由此, 基于(1)式的 MTS 模糊模型与基于(3)式的 NRBF 网络模型是等价的, 即 NRBF 网络中每一项 $\omega_i\Psi(\cdot)$ 对应了一个规则 R^i. 这样, MTS 模型中前件结构的模糊子域划分问题就可转化为寻找 NRBF 网络中心点问题.

2 基于 RBF 网络的学习算法(The learning algorithm based on the RBF network)

在 RBF 网络的框架下, 确定 MTS 模糊模型中规则的个数、前件隶属度函数及后件中线性自回归模型的参数.

给定输入输出数据样本集 $D=\{X(t),y(t),t=1,\cdots,N\}$, 假设 MTS 模型的规则个数为输入样本个数, 则(2)式的矩阵形式为:

$$Y=U\cdot F+\epsilon \quad (5)$$

其中: $Y=[y(1),y(2),\cdots,y(N)]^T$, $F=[f^1,f^2,\cdots,f^N]^T$, $\epsilon=[\epsilon(1),\epsilon(2),\cdots,\epsilon(N)]^T$, $U=[u_1,u_2,\cdots,u_N]$, 而 $u_i=[\mu_i(1),\mu_i(2),\cdots,\mu_i(N)]^T$ 这里 $\mu_i(j)=\mu_{A^i}(X(j))$, $i,j=1,\cdots,N$

由(4)式, 辨识(5)式所表示的 MTS 模型的前件结构问题可转化为基于 RBF 网络表示的模型中心点的确定问题. 同样假设 RBF 中心点的数目为输入的样本数(也即模糊规则个数), 非线性基函数的中心点值也取输入样本的数值. 给定非线性函数 $\Psi(\cdot)$, 令 $P_i(t)=P_i(X(t))=\Psi[X-C_i,)^T\Sigma^{-1}(X-C_i)]$, 则(3)式的矩阵形式为:

$$Y=P\cdot W+\epsilon \quad (6)$$

其中: Y 和 ϵ 同上, $W=[\omega_1,\omega_2,\cdots,\omega_N]^T$

$$P=[p_1,p_2,\cdots,p_N]$$
$$p_i=[p_i(1),\cdots,p_i(N)]^T \quad i=1,\cdots,N$$

则向量 Y 可表示为 $p_1,p_2,\cdots,p_N$ 向量的线性组合, 但需注意的是这 N 个向量 $p_1,p_2,\cdots,p_N$ 仅仅是确定网络模型结构选项的一个范围, 实际确定的模型结构应是从这 N 个向量中选取的一个很小的子集(如 m 个, $m\ll N$), 即在权衡模型结构简单化和高精度拟合要求两方面取得的一种统一. 因此, 为实现以上选项过程, 我们不仅要考虑各输入数据在 n 维空间中的分布状况, 还要考虑它们在输入输出空间中和输出 y 的线性相关性. 为此我们引入以下选项准则:

假设已选入模型的向量为 $p_1,p_2,\cdots,p_k(k\leqslant N)$, 则对剩余的 $\forall p'_j$, p'_j 在由 $p_1,p_2,\cdots p_k$ 组成的超平面上的投影为:

$$u_j=S_k(S_k^TS_k)^{-1}S_k^Tp'_j \quad (7)$$

其中: $S_k=[p_1,p_2,\cdots,p_k]$ 是 $N\times k$ 矩阵.

p'_j 和向量 $p_1,p_2,\cdots,p_k$ 线性无关的部分为: $E_j=p'_j-u_j=[e_j(1),e_j(2),\cdots,e_j(N)]^T$, 则 E_j 表示了 $\forall p'_j$ 和已选入模型的向量所代表的各聚类的贴近程度. 同时我们还应考虑它对拟合输出 Y 的显著性, 计算:

$$J_j=\frac{E_j^TY}{E_j^TE_j}=\left(\sum_{t=1}^{N}e_j(t)y(t)\right)^2\Big/\sum_{t=1}^{N}e_j^2(t) \quad (8)$$

显然 J_j 表示了 p'_j 在拟合 Y 时所做的局部贡献.

为得到精简的模糊模型, 我们使用了 AIC 统计准则对模型进行检验, 如在第 k 步,

$$AIC_k=N\ln\Xi_k^2+2k/N \quad (9)$$

其中: $\Xi_k^2=\frac{1}{N}\sum_{t=1}^{N}[y(t)-\hat{y}_k(t)]^2$, k 是当前选入模型中的项的个数, N 为样本数.

总结以上讨论, 基于逐步回归方法确定模型结构的选项步骤简述如下:

(1) 给定样本输入输出数据 $D=\{X(t),y(t),t=1,\cdots,N\}$ 和非线性函数 $\Psi(\cdot)$ (如取高斯函数). 根据样本数据构造(6)式中各向量 Y,W,p_i,ϵ 且置 k

= 1.

(2) 首先计算每一 $p_i(i=1, \cdots, N)$ 对输出 Y 的相关性, 取其中最大的相应项记为 p'_1, 且置 $k=2$.

(3) 在 $k(k \geqslant 2)$ 时, 对于剩余的 $N-k+1$ 个项 p_j, 基于(7)、(8) 计算出每项的 J_j, 取 $\max_j(J_j)$ 对应的项记为 p'_k. 对于选入 $p'_1, \cdots, p'_k$ 的由最小二乘算法计算拟合输出和相应参数, 由(9) 式的 AIC 准则检验确定是否把 p'_k 选入模型或算法结束.

(4) 同时应考虑在 $k(k \geqslant 2)$ 时, 对于(3) 步中选入模型的 p'_k, 检验其对前选入项的影响, 去掉某一项 $p'_i, i=1, \cdots, k-1$, 重新计算模型的拟合输出和相应参数, 由 AIC 准则检验确定是否应剔除某项 p'_i, $i=1, \cdots, k-1$ 或继续转入(3) 步.

(5) 重复(3)、(4) 步, 根据 AIC 统计准则进行检验直至既无选入项也无剔除项为止.

我们在文献[6]中提出一种基于逐步回归过程的递推算法, 以上(3)、(4) 步中每增加一项或每删除一项使用最小二乘算法拟合输出的矩阵计算均可在前一步的基础上利用[6]中提出的方法递推得到, 且可同时计算出所选入的项, 拟合输出的加权量及相应的拟合误差、统计准则等. 这不仅使整个结构确定过程简单化, 而且大大减少了计算量, 并可在线计算.

在 MTS 模型的前件结构和相应隶属度函数确定后, 则其后件结构的参数估计能很容易得到, 如后件结构是常数, 则由以上辨识出的 RBF 各基函数的加权值 $\omega_i(i=1, 2, \cdots, m)$ 恰可以对应于各规则的后件分量, 只需乘以一规范化常数; 若其后件结构是输入的线性组合形式, 即 $y^i = f^i(X) = b_0^i + \sum_{j=1}^{n} b_j^i x_j$, 则由(2) 式, 可得:

$$y(t) = \sum_{i=1}^{m} \mu_{A^i}(X(t)) f^i(X(t)) = \sum_{i=1}^{m} \mu_{A^i}(X(t))(b_0^i + \sum_{i=1}^{m} b_j^i X_j(t)) \qquad (10)$$

对(10) 式可采用最小二乘算法来估计后件结构中的参数, 这里不再赘述.

3　应用实例(Application example)

为了验证本算法的有效性, 我们进行了实例研究. 在文献[7]中我们曾使用 FCM 算法和遗传算法相结合研究了 1995 年中国宏观经济数据和行业经济数据 9 种重要指标对上交所青岛海尔从 1995, 2 - 1996, 2 的影响, 并建立若干模糊规则, 以辅助决策者进行短期或中期预测. 这 9 种经济指标分别是工业产值、产品销售率、固定资产投资总额、海关出口率、海关进口率、商品零售总额、居民销售价格指数、商品零售价指数和储蓄存款. 为简单起见, 我们只考虑一个月的经济指标并且滞后取作 30, 通过对这九种经济指标的预处理, 可得出相关性最大的经济指标是储蓄存款、工业产值及产品销售率, 这三种指标一个月的增率作为模糊 IF-THEN 规则的输入而将青岛海尔的平均价格波动作为输出, 在文献[7]中, 利用 FCM 算法得出的划分个数分别为 3, 2 和 2, 共 12 个规则. 这里我们基于 MTS 模型和 RBF 网络集成对该实例进行建模. 取 RBF 网络中的非线性变换 $\Psi(\| X(t) - C_i \|) = \exp(-\| X(t) - C_i \|^2 / \beta^2)$ 为高斯函数, 其中 $\beta^2 = 2$. 用本文给出的算法得到的最终 RBF 网络模型是:

$$\begin{aligned} \hat{y}(t) = {} & 3.574606\Psi(\| X(t) - X(5) \| \\ & + 3.897508\Psi(\| X(t) - X(2) \| \\ & + 2.079654\Psi(\| x(t) - X(1) \| + \epsilon(t) \end{aligned}$$

其中: $X(1)$、$X(2)$ 和 $X(5)$ 分别为训练数据中的第一、第二和第五项输入数据. 则相应的 MTS 模型规则个数为 3, 基于(10) 式, 由递推最小二乘法可得到具有线性结构的规则后件的各参数, 这样, 最终的 MTS 模糊模型为:

$$R^1: \text{IF } X \text{ is } A_1, \text{ THEN } y^1(X) = -9.954810 - 7.862705x_1 - 11.481493x_2 + 12.287622x_3$$

$$R^2: \text{IF } X \text{ is } A_2, \text{ THEN } y^2(X) = 4.206620 + 6.720594x_1 - 24.543156x_2 + 7.665077x_3$$

$$R^3: \text{IF } X \text{ is } A_3, \text{ THEN } y^3(X) = 66.532799 - 25.503090x_1 - 10.575146x_2 + 5.592962x_3$$

参　考　文　献(References)

1　T Takagi, M Sugeno. Fuzzy Identification of a Fuzzy Model and its Application to Modeling and Control, IEEE Trans. On System, Man and Cybernetics, 1985, **15**: 116~ 132

2　H Bersini, G Bontempi. Now Comes the Time to Defuzzify Neuro-fuzzy Models, Fuzzy Sets and Systems, 1997, **90**: 161~ 169

3　L Wang, R Langari. Complex Systems Modeling Via Fuzzy Logic, IEEE Trans. On Systems, Man and Cybernetics-part B Cybernetics, 1996, **26**(1)

4　J Q Chen, Y G Xi, Z J Zhang. A Clustering Algorithm for Fuzzy Model Identification, Fuzzy Sets and Systems, 1998, **98**: 319~ 329

5　L Xu, A Krzyzak, A L Yuille. On Radial Basis Function Nets and Kernel Regression: Statistical Consistency, Convergence Rates, and Receptive Field Size, Neural Networks, 1994, **7**(4): 609~ 628

(下转第 239 页)