

文章编号: 1001-0920(2006)09-0984-07

基于协同进化算法的高维模糊分类系统的设计

张永, 邢宗义, 向峥嵘, 胡维礼

(1. 南京理工大学自动化系, 南京 210094)

摘要: 基于协同进化算法, 提出一种高维模糊分类系统的设计方法。首先定义系统的精确性指标, 给出解释性的必要条件, 利用聚类算法辨识初始模型。相互协作的3类种群分别代表系统的特征变量、规则前件和模型隶属函数的参数, 适应度函数采用3类种群合作计算的策略, 在算法运行中利用基于相似性的模型简化技术约简模糊系统, 最后利用该方法对Wine问题进行研究。仿真结果表明该方法能够对高维分类问题的特征变量进行选择, 同时利用较少规则和模糊集合数达到较高的识别率。

关键词: 模糊分类系统; 模糊聚类; 遗传算法; 协同进化算法; 解释性; 精确性

中图分类号: TP273 **文献标识码:** A

Design of High-dimensional Fuzzy Classification Systems Based on Cooperative Coevolutionary Algorithm

ZHANG Yong, XING Zong-yi, XIANG Zheng-rong, HU Wei-li

(Department of Automation, Nanjing University of Science and Technology, Nanjing 210094, China
Correspondent: ZHANG Yong, E-mail: zy69813@gmail.com)

Abstract: An approach to construct high-dimensional fuzzy classification systems is proposed based on multi-objective cooperative coevolutionary algorithm. The precision index is defined, and the necessary conditions of interpretability are analyzed. The initial fuzzy system is identified using fuzzy clustering algorithm. Three cooperative species represent the relevant features, the structure and parameters of the fuzzy systems respectively. The fitness function is calculated on the cooperation of individuals from the three species. The similarity-driven rule based simplification method is applied to reduce the fuzzy model. By computational simulation on the wine classification problem, the proposed approach is able to generate compact fuzzy rule bases with high classification ability.

Key words: Fuzzy classification systems; Fuzzy clustering; Genetic algorithm; Coevolutionary algorithm; Interpretability; Precision

1 引言

如何从数据中自动构造模糊分类系统是近年研究的热点。与神经网络等模型相比, 模糊分类系统的知识表达形式和推理机制符合人类思维习惯, 其结构和隶属函数参数具有明显的物理意义。人们可通过易于理解的模糊规则洞察分类系统的内部运行机理, 即解释性是模糊分类系统最显著的特征, 尤其在医学、金融等领域, 解释性甚至成为构建分类系统时

的首要目标。

随着分类问题维数和复杂性的提高, 利用传统的方法构造模糊分类系统主要存在以下几个问题而使模型不具备解释性: 1) 特征变量的维数存在冗余; 2) 模糊规则数比实际需要的多; 3) 模糊集合的数量与参数设计不合理。所以在保证模型精确性的情况下如何提高模型解释性是当前研究的热点。许多学者进行了相关研究。文献[1~4]给出了模糊模型解

收稿日期: 2005-08-03; 修回日期: 2005-09-16

基金项目: 国家自然科学基金项目(60332020, 60474034)。

作者简介: 张永(1969—), 男, 江苏连云港人, 博士生, 从事模糊建模、神经网络等研究; 胡维礼(1941—), 男, 江苏南通人, 教授, 博士生导师, 从事智能控制与智能系统等研究。

释性的一些必要条件。文献[5~9]给出了提高模糊模型解释性的一些具体方法。遗传算法在众多方法中因其简单通用、鲁棒性强、适于并行处理而得到广泛的应用,目前,匹茨堡方法^[10]较为常用。然而分类问题维数和复杂性的提高,必然导致该类方法中染色体位数变长,算法的搜索空间变大,会引起遗传算法的未成熟收敛,对问题的规模比较敏感,收敛速度慢等缺陷。

为解决遗传算法的上述问题,文献[11]提出了协同进化的思想;文献[12]发展并提出协同进化的一般模型,利用不同种群之间的相互适应,共同进化,合作产生问题的完整解;文献[13]把协同进化算法应用于一类模糊分类系统的建模研究,用两类种群的协同进化,对特征变量没有进行选择,不涉及高维分类系统的建模问题。

本文提出一种基于多目标协同进化算法(MOCOEA)的高维模糊分类系统的设计方法。首先定义系统的精确性指标,给出解释性的必要条件;利用模糊聚类算法辨识初始的模糊模型;相互协作的3类种群采用不同的编码方式,分别代表系统特征变量、模型的规则前件和模型隶属函数的参数;适应度函数考虑模糊分类系统的精确性指标和解释性指标,采用3类种群合作计算的策略;在协同进化算法运行中同时利用基于相似性的模型简化技术约简模糊分类系统,有效地保证了模糊分类系统的解释性;最后利用该方法对wine分类问题进行研究。仿真结果同其他文献结果进行比较,表明该方法能对高维模糊分类系统的特征变量进行选择,同时利用较少的模糊分类规则和模糊集合数达到较高的识别率。

2 预备知识

2.1 模糊分类系统

考虑 n 维 N_c 类 N 样本的分类问题,其中 $x \in X \subseteq R^n, x = (x_1, x_2, \dots, x_n)$ 为特征变量, $\{g_1, g_2, \dots, g_M\}$ 为输出类别,则模糊分类系统典型规则形式为

R_i : if x_1 is A_{i1} and x_2 is A_{i2} and ... and x_n is A_{in} ,
then the pattern (x_1, \dots, x_n) belongs to g_i .

其中 A_{i1}, \dots, A_{in} 为定义在输入论域中的隶属函数,可取三角形、高斯型、梯形等。本文采用高斯型函数

$$A_{ij}(x_j) = \exp\left(-\frac{(x_j - v_{ij})^2}{2\sigma_{ij}^2}\right), \quad (1)$$

其中 v_{ij} 和 σ_{ij} 分别代表函数的中心和方差。

对未知样本 x_k ,模糊分类系统的输出采用“胜者为王”的策略,即系统的输出为具备最大激励强度的规则所对应的类别输出

$$x_k \quad g^{i^*}, i^* = \arg \max (\beta_i(x_k)). \quad (2)$$

其中 β_i 为第 i 条规则的激励强度

$$\beta_i(x_k) = \prod_{j=1}^n A_{ij}(x_{jk}). \quad (3)$$

2.2 精确性与解释性

给定样本变量 x_k ,其分类误差定义为

$$e_k = \begin{cases} 0, & \text{如果 } x_k \text{ 被正确分类;} \\ 1, & \text{如果 } x_k \text{ 被错误分类} \end{cases} \quad (4)$$

则定义衡量模糊分类系统的精确性的指标为

$$J_{ERR} = \frac{1}{N} \sum_{k=1}^N e_k \quad (5)$$

与精确性等可以量化的性能指标不同,模糊分类系统的解释性目前尚无明确的标准和定义,但一般认为模糊系统的解释性与模型结构、特征变量和模糊规则数目、隶属函数特性等密切相关,现将主要因素陈述如下^[1~4]:

1) 特征变量数目:人们很难通过高维模糊系统来分析系统行为,因此为了提高模型的解释性需进行特征变量的选择,使模糊系统采用尽可能少的特征变量。

2) 模糊规则数目:模糊系统的规则数目越多,其解释性越低。经验认为,可解释的模糊系统其规则数目不超过10个,这是由人在理解、推理时的思维能力所决定的。

3) 隶属函数特性:隶属函数必须是凸的,其划分必须是完备的,即对于任何特征变量,在其论域内的任何值,至少有一个隶属函数相对应,在形式上表现为隶属函数之间存在位置的交叉。隶属函数划分必须是可区分的,即对于同一变量,隶属函数之间存在明显的位置区别,以便赋予一定的语义项。

4) 模糊规则库的完整性、一致性和精简性:模糊规则要完整覆盖特征论域,对每一个有效的特征变量组合,至少有一条模糊规则被激励,即完整性;模糊规则之间必须相容而不能有任何两条规则相互矛盾,即一致性;在规则数目尽可能小的前提下,不能包含冗余规则,如某规则的前件是另一规则的子集等,即精简性。

3 模糊分类系统的初始化

首先利用模糊聚类算法辨识初始模糊模型,保证协同进化算法从较合理的初始模型开始学习,这样可以有效节省搜索的空间和时间。本文采用Gustafson-Kessel(GK)聚类算法离线辨识模糊系统的前件参数^[14]。

为确定规则后件类的标号值,定义如下函数:

$$M_{ij} = \frac{\sum_{k=1}^N u_{ik} f_j(x_k)}{\sum_{k=1}^N f_j(x_k)}, \quad (6)$$

其中 μ_{ik} 是第 k 个数据相对于第 i 个聚类中心的隶属度,且满足以下条件:

$$\mu_{ik} \in [0, 1], \quad (7)$$

$$\sum_{i=1}^C \mu_{ik} = 1, 0 < \sum_{k=1}^N \mu_{ik} < N, \quad (8)$$

$$f_j(x_k) = \begin{cases} 1, & \text{如果 } x_k \text{ 属于类 } j; \\ 0, & \text{如果 } x_k \text{ 不属于类 } j. \end{cases} \quad (9)$$

则对于第 i 条规则,其后件类的标号值 g_i^* 为

$$i^* = \arg(\max(M_{ij})), j = 1, 2, \dots, N_a \quad (10)$$

4 基于相似性的模糊模型简化

4.1 模糊集合的相似性分析与融合

模糊模型的隶属函数存在 3 种类型的冗余:第 1 种是两个模糊集合相似,重叠区域过大,这是最常见的模糊集合冗余形式;第 2 种是隶属函数以较大值覆盖整个论域;第 3 种是隶属函数接近于单点集合

第 2 种和第 3 种的隶属函数冗余,由于其不存在解释性的实际意义,在满足精度的前提下,一般在对应的规则前件中直接去除.对于第 1 种冗余,本文采用相似性测度来评判两个隶属函数的相似性.对模糊集合 A 和 B ,定义相似性测度如下^[8]:

$$S(A, B) = \frac{\sum_{k=1}^N [\min(\mu_A(x_k), \mu_B(x_k))]}{\sum_{k=1}^N [\max(\mu_A(x_k), \mu_B(x_k))]} \quad (11)$$

其中: \min 和 \max 分别为最小和最大算子; S 为定义在 $[0, 1]$ 间的相似性测度, $S = 1$ 表示两个集合完全相等, $S = 0$ 表示两个集合没有交叉或重叠

如果两个模糊集合 A 和 B 的相似性测度大于预先设定的阈值,那么集合 A 和 B 可以融合为新的集合 C .对于本文所采用的高斯型隶属函数,由集合 A 和 B 融合生成的新集合 C 的参数如下:

$$\begin{cases} v_c = (v_A + v_B)/2, \\ \sigma_c^2 = (\sigma_A^2 + \sigma_B^2)/2 \end{cases} \quad (12)$$

阈值的大小直接影响模糊系统的性能,阈值越小,得到的模糊系统精度越低而解释性越高,一般阈值取 $[0.4, 0.7]$,具体可根据实际系统要求选择

模糊集合融合过程需要反复迭代进行.在每一次迭代过程中,对每一个变量的所有模糊集合进行两两相似性分析,相似性测度大于阈值的两个模糊集合融合为新的集合.迭代反复进行,直到没有任何两个模糊集合的相似性测度大于阈值,然后再将第 2 类和第 3 类模糊集合删除,从而完成整个模糊集合的相似性融合过程

4.2 模糊规则的相似性分析与融合

在模糊规则的融合过程中,会出现两种类型的规则融合:1) 两条规则中相关的模糊集合被融合,

或者在模型结构种群进化过程中出现两条规则编码相同,导致两条模糊规则前件完全相同.此时可以通过规则的相似度^[7]来计算规则之间的相似性.考虑如下两条规则:

R_i: if x_1 is $A_{i1}(x_1)$ and x_2 is $A_{i2}(x_2)$ and ... and x_n is $A_{in}(x_n)$,

then the pattern (x_1, \dots, x_n) belongs to g_i ;

R_j: if x_1 is $A_{j1}(x_1)$ and x_2 is $A_{j2}(x_2)$ and ... and x_n is $A_{jn}(x_n)$,

then the pattern (x_1, \dots, x_n) belongs to g_j .

这两条规则前件的相似度 SRA 定义如下:

$$SRA(i, j) = \frac{1}{n} \sum_{k=1}^n S(A_{ik}, A_{jk}). \quad (13)$$

其中 $S(\bullet)$ 的计算如式(11)所示,当该值为 1 时,表明两条规则的前件相同;当该值较小时,表明两条规则的前件有很大的不同;对于任意一组模糊规则,若其相似度 SRA 大于给定阈值 $[0.9, 1]$,可将该组规则随机删除一条,实现规则的约简.2) 模型结构种群在进化过程中会出现个体中某一条规则的编码全为零的情况,代表该规则中没有相关变量的模糊集合被激励,这时该个体反编码为模糊模型时可删除该条规则.因为模糊分类系统的规则的后件为类标,故不考虑规则的后件的相似性

5 多目标协同进化算法(MOCOEA)

协同进化算法采用分而治之的思想将模糊分类系统分解为相互合作的 3 类种群:第 1 类种群描述模糊分类系统特征变量;第 2 类种群描述模型的规则前件;第 3 类种群描述模型隶属函数的参数.3 类种群分别采用不同的遗传算法进化,适应度函数考虑模糊分类系统的精确性指标和解释性指标,采用 3 类种群合作计算的策略

5.1 模型的分解和各种群染色体编码

对于式(1)所示的模糊分类系统,在能保证合理解决系统的精确性和解释性的前提下,分解的种群越少,协同进化算法的计算越简单,搜索空间和时间越少.本文将模糊模型分解为以下 3 个种群:特征变量种群、规则前件种群和隶属函数参数种群

5.1.1 特征变量种群

对于模糊分类系统特征变量种群,采用二进制编码方式.其编码形式如图 1 所示

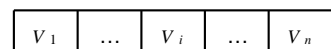


图 1 二进制编码形式

这里 V_i 取值为 $\{1, 0\}$,当 V_i 取值为 1 时,表示第 i 个特征变量被选择;当 V_i 取值为 0 时,表示第 i 个特征变量被屏蔽

给定种群大小为 L , 第一条染色体取值为全“1”, 表示选择全部特征变量, 其余的 $L - 1$ 个染色体在二进制空间内随机均匀生成, 从而形成初始特征变量种群

5.1.2 规则前件种群

定义模型规则前件包括: 模型的规则数, 规则涉及到的变量所对应的相关模糊集合. 采用二进制编码, 使用定长染色体

首先随机产生初始结构种群, 其中二进制染色体反编码后的表现型如图 2 所示

{ 规则 1 }	{ 规则 i }	{ 规则 c }
A_{11} ... A_{1n}	A_{ij} ... A_{ci}	A_{cn}

图 2 二进制染色体反编码后的表现型

这里 A_{ij} 取值为 $\{0, 1, 2, \dots, c\}$, c 为初始聚类数. 当 A_{ij} 取值为 0 时, 表示第 i 条规则中的第 j 个变量为变量无关项^[9], 这是一条不完整规则; 当 A_{ij} 取其他值时, 则第 i 条规则中的第 j 个变量所对应的相关模糊集合被激励

给定种群大小为 L , 在二进制空间内随机均匀生成初始模型规则前件种群

5.1.3 隶属函数参数种群

模糊聚类产生的初始模型的前件参数, 可直接编码产生第一条染色体, 待编码的参数为隶属函数的中心 v_{ij} 和方差 σ_{ij} , 因此每条染色体共有 $2 \times c \times n$ 个实数. 对于第一条染色体其编码为

$$H_1 = (v_{11}, \dots, v_{cn}, \sigma_{11}, \dots, \sigma_{cn}). \quad (14)$$

给定种群大小为 L , 染色体表示为 $H_p, p = 1, 2, \dots, L$, 给定搜索空间 $[H^{\min}, H^{\max}]$ 为

$$H^{\min} = (v_{11}^{\min}, \dots, v_{cn}^{\min}, \sigma_{11}^{\min}, \dots, \sigma_{cn}^{\min}), \quad (15)$$

$$H^{\max} = (v_{11}^{\max}, \dots, v_{cn}^{\max}, \sigma_{11}^{\max}, \dots, \sigma_{cn}^{\max}), \quad (16)$$

其中 $v_{ij}^{\max}, v_{ij}^{\min}, \sigma_{ij}^{\min}, \sigma_{ij}^{\max}$ 为对应隶属函数的中心和方差的最大最小约束值. 则其余的 $L - 1$ 个染色体以染色体 H_1 为中心, 在搜索空间内随机均匀生成, 从而形成初始参数种群

5.1.4 子种群染色体之间的合作

3 类种群中的染色体的合并经过反编码后才能产生完整的模糊分类系统, 且染色体的合并有顺序要求, 为了更清楚地说明子种群染色体之间如何相互合作形成模糊分类系统, 给出下面的例子.

设规则数 $c = 5$, 特征变量的维数 $n = 8$, 从 3 类种群中随机选取 3 条染色体进行合并, 如图 3 所示

图 3 中, 第 1 行是特征变量染色体编码; 第 2 行是模型规则前件的表现型; 第 3 行是隶属函数参数编码, 则上述编码表示的模型规则前件为

1	0	0	1	0	1	0	0
02304011	10341400	21135521	01235121	11545411			
v_{11} ... v_{18}	v_{21} ... v_{28}	v_{31} ... v_{38}	σ_{11} ... σ_{18}	σ_{21} ... σ_{28}	σ_{31} ... σ_{38}		

图 3 染色体之间的合并

R_1 :

R_2 : if x_1 is $A_{21}(v_{11}, \sigma_{11})$ and x_4 is $A_{24}(v_{44}, \sigma_{44})$
and x_6 is $A_{26}(v_{46}, \sigma_{46})$,

R_3 : if x_1 is $A_{31}(v_{21}, \sigma_{21})$ and x_4 is $A_{34}(v_{34}, \sigma_{34})$
and x_6 is $A_{36}(v_{56}, \sigma_{56})$,

R_4 : if x_4 is $A_{44}(v_{34}, \sigma_{34})$ and x_6 is $A_{46}(v_{16}, \sigma_{16})$,

R_5 : if x_1 is $A_{51}(v_{11}, \sigma_{11})$ and x_4 is $A_{54}(v_{44}, \sigma_{44})$
and x_6 is $A_{56}(v_{46}, \sigma_{46})$,

其中: 若 R_1 中选择的特征变量所对应的模糊集合均没有被激励, 则在基于相似性的模糊规则融合过程中被删除; R_4 中 x_1 为变量无关项; R_2 和 R_5 形式上不同, 但其模糊集合参数完全相同, 则在基于相似性的模糊规则融合过程中被随机删除一条. 所以以上编码实际只表示含有 R_2, R_3 和 R_4 的 3 条规则的模糊模型, 相关的模糊集合数为 7 个, 若此模糊分类系统的精确性指标能达到要求, 则该 3 条染色体代表的是一个解释性好、精确性高的模糊模型

通过以上步骤获得模糊模型的规则前件, 使用式 (7) 和 (10) 计算模糊分类系统的后件类标, 所以后件类标在协同进化算法的 3 类种群中没有体现

5.2 种群个体的适应度计算策略

不同于常规的遗传算法, 在协同进化中, 个体适应度的计算依赖于该个体与其他种群的代表合作产生该个体的适应度的值

种群代表可以是种群全体或者典型个体, 考虑到计算的复杂性, 本文选用典型个体作为种群代表. 每个种群的代表选取由该种群的两类个体构成: 一类是抽取 N_{cf} 个适应度最好的个体; 另一类是随机抽取 N_{cr} 个个体. 则该种群代表中个体数为 $N_{rep} = N_{cf} + N_{cr}$.

个体适应度的计算方法为: 该个体与其他种群代表合作产生 N_{rep} 个模糊模型, 对每个模型计算其适应度, 取其中最优的适应度作为该个体的适应度值. 种群代表的产生和个体适应度的计算如图 4 所示. 协同进化算法对个体适应度的计算经历了一个局部学习的过程, 体现了协同进化与其他遗传算法的不同, 强调了个体进化与学习的交互作用

5.3 适应度函数

种群中的个体与其他种群代表中的个体合作, 经过反编码产生模糊模型的前件, 首先利用基于相似性的模型简化技术约简模糊模型, 保证模型的解

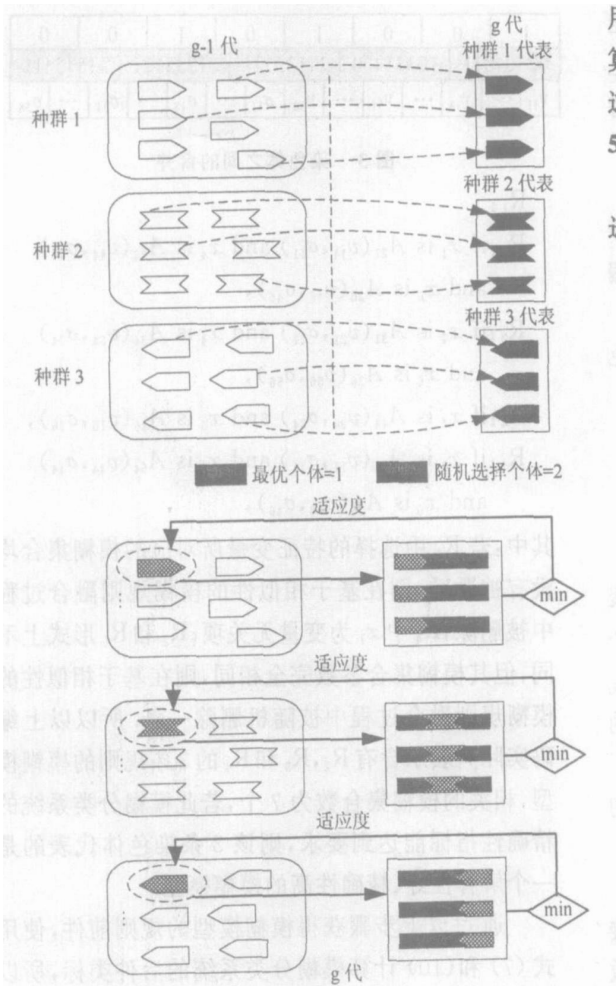


图4 种群代表的产生和个体适应度的计算策略

释性 在设计个体的适应度函数时,模型的精确性指标考虑 J_{ERR} , 在解释性指标上主要考虑特征变量数目和规则库中模糊集合总数分别为

$$J_{nVar} = \sum_{i=1}^n V_i^a, \quad (17)$$

$$J_{nFS} = \sum_{j=1}^{J_{nVar}} S_j^a, \quad (18)$$

其中: V_i^a 为 1 时表示第 i 个特征变量被选择, S_j^a 为被选择特征变量 x_j 中被激励的模糊集合数, n 为特征变量维数 3 个指标函数均越小越好, 采用加权适应度函数为

$$Fit = w_1 J_{ERR} + w_2 J_{nVar} + w_3 J_{nFS} \quad (19)$$

其中加权因子 w_1, w_2, w_3 为正实数, 在计算适应度函数时预先根据经验给定

5.4 种群的遗传操作

MOCOEA 算法将模糊分类系统分解为相互合作的 3 类种群, 采用不同编码方式, 但都利用遗传算法进化 遗传操作主要包括: 选择、交叉和变异, 3 类种群都选用精英保留和随机遍历抽样法相结合的选择操作 由于 MOCOEA 算法采用分而治之的思想将复杂的问题分解为简单的子问题, 且算法中又采

用精英保留策略, 为防止算法未成熟收敛, 协同进化算法的交叉概率和变异概率要比常规遗传算法高, 这是协同进化算法的特点

5.5 MOCOEA 算法步骤

利用 3 类不同的遗传算法控制 3 类种群的协同进化, MOCOEA 算法的伪代码如下:

```

Begin MOCOEA
g := 0
for each species S
    Initialize population Pops(0)
    Evaluate population Pops(0)
    Reps(1) =
    Select - representation[Pops(0)]
end for
while not done
    for each species S
        g := g + 1
        Es(g) = Elite - select[Pops(g - 1)]
        Pops1(g) = Select[Pops(g - 1)]
        Pops2(g) = Crossover[Pops1(g)]
        Pops3(g) = Mutate[Pops2(g)]
        Pops(g) = Pops3(g) + Es(g)
        Models1(g) =
        Merge[Pops(g), Reps(g)]
        Models2(g) =
        Similarity - simplify[Models1(g)]
        Models(g) =
        Models2(g) + Class[Models2(g)]
        Evaluate Fitness Models(g)
        Reps(g + 1) =
        Select - representation[Pops(g)]
    end for
end while
end MOCOEA
    
```

其中: $Rep_s(g)$ 和 $Rep_{-s}(g)$ 表示另 2 类种群代表; $Model_s^1(g)$ 表示 3 类种群中染色体合并后代表的模糊模型前件(含有冗余信息); $Model_s^2(g)$ 表示经过基于相似性模型简化技术约简后的模糊模型前件; $Class[Model_s^2(g)]$ 表示利用式(7)和式(10)计算模糊分类系统的后件类标; $Model_s(g)$ 表示完整的模糊分类系统; $Evaluate\ Fitness\ Model_s(g)$ 计算策略如图 4 所示, 计算结果作为 $Pop_s(g)$ 相应染色体的适应度

6 仿 真

本文利用典型 benchmark 问题 Wine 分类问题

检验MOCOEA 算法自动构造精确性和解释性模糊分类系统的性能, 仿真程序在Matlab 7.0 中运行完成.MOCOEA 算法各参数设置如表1 所示

表1 MOCOEA 算法参数设置

参 数	值	参 数	值
最大进化代数	100	种群3 交叉率	0.85
种群初始个体数	40	种群3 变异率	0.1
种群1,2 交叉率	1	种群3 精英率	0.025
种群1,2 变异率	0.2	N_{cf}	1
种群1 精英率	0.05	N_{cr}	2
种群2 精英率	0.1	权值 $w_1 = w_2 = w_3$	1

Wine 分类问题的分类目标为: 正确分类来自意大利同一地区 3 类不同的培育植物酿制的酒, 其中第1 类59 例, 第2 类71 例, 第3 类48 例 13 维特征变量分别为: x_1 alcohol, x_2 malic acid, x_3 ash, x_4 alkalinity of ash, x_5 m agnesium, x_6 total phenols, x_7 flavonoids, x_8 nonflavonoid phenols, x_9 proanthocyanins, x_{10} color intensity, x_{11} hue, x_{12} OD 280/OD 315 of diluted wines, x_{13} pro line

初始聚类数分别设定为 3, 4, 5 和 6, 进行 4 次协同进化计算, 仿真结果如表 2 所示 在MOCOEA 算法运行中, 模糊集合相似性融合的阈值设为 0.4, 模糊集合相似于整个论域融合的阈值设为 0.8, 模糊规则相似性融合阈值设为 1.

表2 Wine 分类问题仿真结果

初始聚类数	模糊规则数	模糊集合总数	相关特征变量	正确分类率/%
3	3	5	$x_1 x_3 x_7 x_{12}$	96.06
4	3	4	$x_1 x_4 x_7 x_8$	92.69
5	5	7	$x_1 x_7 x_{12} x_{13}$	97.19
6	3	4	$x_1 x_{12}$	92.12

如表2 所示, 聚类数为5 时所获得模糊分类系统准确率最高 图5 为此时系统的隶属函数分布 表3 为该模糊分类系统的规则库

表3 Wine 模糊分类系统规则库

R1:	if x_1 is high and x_7 is medium and x_{13} is high then output is class1
R2:	if x_7 is medium and x_{12} is low and x_{13} is low then output is class2
R3:	if x_1 is high and x_7 is medium and x_{12} is medium and x_{13} is high then output is class1
R4:	if x_1 is low and x_7 is medium and x_{12} is medium and x_{13} is low then output is class2
R5:	if x_1 is high and x_{12} is low and x_{13} is low then output is class3

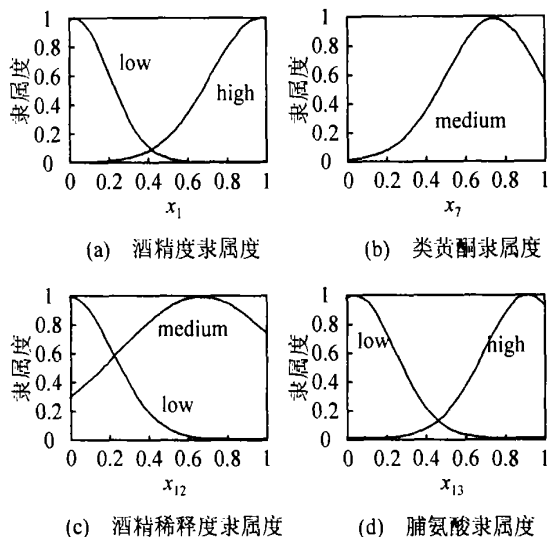


图5 Wine 模糊分类系统属函数分布

表4 给出本文方法同其他文献建模方法的性能比较 文献[9, 15]虽然有较高的分类率, 但其规则数为 60, 模型基本没有解释性; 文献[1]方法的分类率为 60, 模型基本没有解释性; 文献[1]方法的分类率比本文略高, 但其模糊集合总数是本文方法的 3 倍

表4 Wine 分类问题不同建模方法的性能比较

	规则数	集合数	正确分类率/%
文献[15]	60	—	100
文献[1]	3	21	98.31
文献[9]	60	—	99.4
本文方法	5	7	97.19

7 结 语

本文提出一种基于多目标协同进化算法的高维模糊分类系统的设计方法, 并论述了与之相关的问题 该方法能对高维模糊分类系统的特征变量进行选择, 同时利用较少的模糊分类规则和模糊集合数达到较高的识别率

对于本文方法, 还存在以下问题: 1) 需要多次设置 初始聚类数和多次运行MOCOEA 算法; 2) 适应度函数的权值需根据经验给定且算法运行一次只能获得一个解 解决上述问题, 使算法运行一次获得一组Pareto 最优解将是下一步研究的课题

参考文献 (References)

[1] Setnes M, Roubos H. GA-fuzzy Modeling and Classification: Complexity and Performance [J]. *IEEE Trans on Fuzzy Systems*, 2000, 8: 509-522

[2] Jin Y. Fuzzy Modeling of High-dimensional Systems Complexity Reduction and Interpretability Improvement [J]. *Fuzzy Sets and Systems*, 2000, 8(2): 212-221

[3] Casillas J, Cordón O, Herrera F, et al. Interpretability Improvements to Find the Balance Interpretability-

- accuracy in Fuzzy Modeling: An Overview [A] *Chapter of Interpretability Issues in Fuzzy Modeling* [C] Springer, 2003: 3-22
- [4] Jin Y. *Advanced Fuzzy Systems Design and Applications* [M] New York: Physical-verl, 2003
- [5] Kaynak O, Jezernik K, Szeghegyi A. Complexity Reduction of Rule Based Models: A Survey [A] *Proc of IEEE Int Conf on Fuzzy Systems* [C] Honolulu, 2002: 1216-1222
- [6] Guillaume S. Designing Fuzzy Inference Systems from Data: An Interpretability-oriented Review [J] *IEEE Trans on Fuzzy Systems*, 2001, 9(3): 426-443
- [7] Jin Y, Seelen von W, Sendhoff B. An Approach to Rule-based Knowledge Extraction [A] *Proc IEEE Conf Fuzzy System* [C] Alaska, 1998: 1188-1193
- [8] Setnes M, Babuska R, Kaymak U, et al Similarity Measures in Fuzzy Rule Base Simplification [J] *IEEE Trans on Systems Man and Cybernetics*, 1998, 28(3): 376-386
- [9] Ishibushi H, Nakashima T, Murata T. Performance Evaluation of Fuzzy Classifier Systems for Multidimensional Pattern Classification Problems [J] *IEEE Trans on Systems, Man, and Cybernetics — Part B: Cybernetics*, 1999, 29(5): 601-618
- [10] Michalewicz Z. *Genetic Algorithms + Data Structures = Evolution Programs* [M] 3rd ed Heidelberg: Springer-Verlag, 1996
- [11] Paredis J. *Coevolutionary Computation, Artificial Life 2* [M] MIT Press, 1995: 355-375
- [12] Potter M A, De J K A. Cooperative Coevolution: An Architecture for Evolving Coadapted Subcomponents [J] *Evolutionary Computation*, 2000, 8(1): 1-29
- [13] Carlos Andres Pena-Reyes, Moshe Sipper. Fuzzy CoCo: A Cooperative Coevolutionary Approach to Fuzzy Modeling [J] *IEEE Trans on Fuzzy Systems*, 2001, 9(5): 727-737.
- [14] Gustafson D, Kessel W. Fuzzy Clustering with a Fuzzy Covariance Matrix [A] *Proc of IEEE Conf on Decision and Control* [C] San Diego, 1979: 761-766
- [15] Corcoran A L, Sen S. Using Real-valued Genetic Algorithms to Evolve Rule Sets for Classification [A] *Proc 1st IEEE Conf on Evolutionary Computation* [C] Orlando, 1994: 120-124

(上接第983页)

参考文献(References)

- [1] Park H, Hong Y, Chang S Y. An Efficient Scheduling Algorithm for the Hot Coil Making in the Steel Mill [J] *Production Planning and Control*, 2002, 13(3): 298-306
- [2] Tang L X, Liu J Y, Rong A Y, et al A Multiple Traveling Salesman Problem Model for Hot Rolling Scheduling in Shanghai Boshan Iron and Steel Complex [J] *European J of Operational Research*, 2000, 124(2): 267-282
- [3] 李耀华, 王伟, 徐乐江, 等 热轧生产轧制计划模型与算法研究 [J] *控制与决策*, 2005, 20(3): 275-279
(Li Y H, Wang W, Xu L J, et al Rolling Plan Model and Algorithm in Hot Rolling Plant [J] *Control and Decision*, 2005, 20(3): 275-279)
- [4] Peter C, Wafa R. Integration of Continuous Caster and Hot Strip Mill Planning for Steel Production [J] *J of Scheduling*, 2000, 3(4): 185-208
- [5] 朱宝琳, 于海斌 冶铸轧生产调度模型及算法研究 [J] *计算机集成制造系统*, 2003, 9(1): 33-36
(Zhu B L, Yu H B. Production Scheduling Model and Algorithm for Steel Making-continuous Casting-hot Rolling Processes [J] *Computer Integrated Manufacturing Systems*, 2003, 9(1): 33-36)
- [6] 吕志民, 徐金梧 一种适用于热送热装生产计划优化的方法 [J] *北京科技大学学报*, 2002, 24(6): 675-678
(Lv Z M, Xu J W. Optimization Method for Hot Charge Rolling Manufacture Plan [J] *J of University of Science and Technology Beijing*, 2002, 24(6): 675-678)
- [7] 郑秉霖, 胡琨元, 常春光 一体化钢铁生产计划系统的研究现状与展望 [J] *控制工程*, 2003, 10(1): 6-10
(Zheng B L, Hu K Y, Chang C G. Status and Expectation of Research on Integrated Planning for Steel Production [J] *Control Engineering of China*, 2003, 10(1): 6-10)
- [8] Marc Gravel, Wilson L Price, Caroline Gagne. Scheduling Continuous Casting of Aluminum Using a Multiple Objective Ant Colony Optimization Metaheuristic [J] *European J of Operational Research*, 2002, 143(1): 218-229
- [9] Merkle D, Middendorf M, Schneck H. Ant Colony Optimization for Resource-constrained Project Scheduling [J] *IEEE Trans on Evolutionary Computation*, 2002, 6(4): 333-346