

基于Q学习算法的变论域模糊控制新算法

余 涛¹, 于文俊², 李章文¹

(1. 华南理工大学 电力学院, 广东 广州 510640; 2. 湖南电力公司 永州电业局, 湖南 永州 425000)

摘要: 变论域模糊控制器的控制函数被“复制”到后代中, 往往存在着“失真”现象, 这种现象的后果是造成算法本身的误差. 针对这一问题, 本文提出了一种基于Q学习算法的变论域模糊控制优化设计方法. 本算法在变论域模糊控制算法基础上提出了一种利用伸缩因子、等比因子相互协调来调整论域的构想, 且通过用Q学习算法来寻优参数使控制器性能指标最小, 使其在控制过程中能够降低“失真率”, 从而进一步提高控制器性能. 最后, 把算法运用于一个二阶系统与非最小相位系统, 实验表明, 该算法不但具有很好的鲁棒性及动态性能, 且与变论域模糊控制器比较起来, 其控制性能也更加提高.

关键词: 变论域模糊控制; Q学习算法; 伸缩因子; 等比因子

中图分类号: TM71 文献标识码: A

A novel optimal method of variable-universe fuzzy control based on Q-learning algorithms

YU Tao¹, YU Wen-jun², LI Zhang-wen¹

(1. College of Electrical Power, South China University of Technology, Guangzhou Guangdong 510640, China;

2. Yongzhou Electric Power Bureau of Electric Power of Hunan, Yongzhou Hunan 425000, China)

Abstract: When the control function of the variable-universe fuzzy controller is “copied” to the offspring, there usually exist some “distortion” phenomena which lead to the error of the algorithms. To deal with this problem, we present a novel optimal method of variable-universe fuzzy control based on Q-learning algorithms. This method adjusts the universe by the contraction-expansion factor and geometric proportional factors, and optimizes the parameters through Q-learning algorithms to minimize the performance index of the controller for reducing the “distortion rate” in the control process and improve the control performance. This method has been applied to a second-order linear system with non-minimum phase, resulting in desirable robustness and dynamic performance. The control performance is even better than that of the variable universe fuzzy controller.

Key words: variable-universe fuzzy control; Q-learning algorithms; contraction-expansion factor; geometric proportional factors

1 引言(Introduction)

模糊控制作为智能控制的一个重要分支, 凭借其无需精确数学模型、鲁棒性强、实时性好等优点而广泛被用于非线性、时变滞后系统中. 然而, 常规模糊控制器存在着稳态精度低、存在调节死区等不足而并不适合用于高精度问题控制^[1~3]. 为了解决以上缺陷, 比较容易想到的一个办法是增加模糊规则的数量, 然而通过增加模糊规则提高控制器稳态的精度, 其代价是降低了控制器的决策速度. 如何提高模糊控制器控制精度而又不影响控制器动态性能近年来已经变成了一个全新的课题, 其中我国学者李洪兴教授从改变隶属度函数论域角度给出了一条全新的思路^[1~3].

文[1]首先指出了模糊控制器的本质是插值控制器. 文[2]在文[1]的基础上认为模糊推理前件的模糊集恰为插值的基函数, 要减少控制稳态误差, 就需要减小模糊集峰点之间的距离. 在文[2]的结论上文[3]进一步提出了变论域模糊控制(variable universe fuzzy control, VUFC)思想, 即在保持模糊规则不变的前提下, 论域随着误差的变小而收缩(亦可随着误差变大而伸扩展). 理论上这种控制器应该有着良好的决策与精度, 可实际上由于控制函数在遗传到后代中存在着“失真”问题, 尽管在控制精度上还是有较大提高, 但作者在研究中发现: 该方法无法保证控制品质的最优性, 且在模糊规则确定好的情况下, 变论域控制器的输入输出伸缩因子的

调节“压力”过大.文[4]首次提出了“失真率”的概念,且把变论域模糊控制系统模拟成一条Markov链,提出了一种计算失真率的方法,可并没有给出其降低失真率的优化方法.文[5]则提出了一种等比因子来调节隶属度函数输入输出论域,可优化方法设计上是固定等比因子,而优化量化因子以及比例因子.

在前人研究^[1~5]的基础上,本文对变论域模糊控制算法进行了进一步改进,通过引进具有在线学习和动态寻优能力的Q学习算法^[6]形成了一种新颖的Q学习-变论域模糊控制算法(variable universe fuzzy control based on Q-learning, Q-VUFC).本文方法可解决以下两方面的问题:1) 弥补变论域模糊控制器的缺点,即能够使“失真”问题得到优化,从而降低由于“失真”所带来控制影响;2) 使控制器保持好的决策速度与控制精度.

2 问题描述(Problem description)

变论域的控制函数的变形,初始控制函数的主要特征随着时间的推移被缩小(或放大)“复制”到“后代”(如 t_1 代, t_2 代)控制函数之中,如图1,初始控制函数来源于初始控制规则,而初始控制规则恰恰是领域专家(如现场操作人员等)经验总结,在“复制”遗传中往往产生“失真”^[2].失真就是指随着论域的误差的减少,初始控制函数并不能完全保持,而产生变形,使得控制无法达到预期的效果.失真的严重后果是对控制器本身的性能的影响.

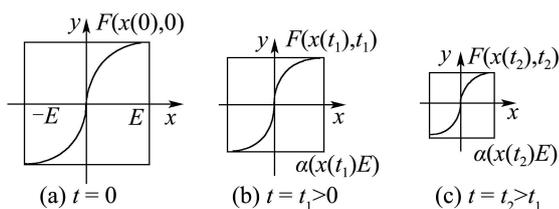


图1 控制器函数随时间变化图

Fig. 1 Control function at different times

这个问题在实际中也是不可避免的,因为采用伸缩因子后,隶属度函数发生变化,对应的输入输出也必然变化^[4],当某一步的控制函数发生畸变时,将直接影响了整个“后代”,这样整个控制器的控制性能就会发生改变,这是一个非常严重的问题,但是这种“问题”严重性可以缓解,也就是说,其“失真率”可以得到降低.失真率就是指一个变论域自适应模糊控制器中,任何一个输入变量在所有可能取值的集合中,使该模糊控制器不能达到预期反馈的概率^[4].

问题的本质在于隶属度函数发生了变化,因此在优化“失真”问题的时候,还得从隶属度函数入手,可以引进另外一个因子,该因子的作用主要能够根据外部信息(如误差、ITAE等),配合伸缩因子而调节隶属度函数参数,使隶属度函数最大限度的符合控

制器的需要,以此来降低“失真率”,维持控制器的良好的性能.

3 控制算法(Control algorithm)

3.1 变论域模糊控制器原理(VUFC algorithm)

给定模糊控制器, $X_i = [-E_i, E_i]$ 是输入变量 $x_i (i=1, 2, 3, \dots, n)$ 的论域,而 $Z = [-U, U]$ 是输出变量 z 的论域. $A = A_{ij} (i=1, 2, 3, \dots, n, j=1, 2, 3, \dots, m)$ 是论域 X_i 上的模糊划分,而 $B = \{B_j\}$ 是论域 Z 上的一个模糊划分, A, B 是语言变量,可以形成如下模糊控制规则:

If x_1 is A_{1j} , and x_2 is A_{2j} , and \dots , and x_n is A_{nj} , then z is $B_j (j=1, 2, \dots, m)$.

设 X_{ij} 为 A_{ij} 的峰点, Z_j 为 B_j 的峰点($i=1, 2, 3, \dots, n, j=1, 2, 3, \dots, m$),模糊控制系统可以表示成一个 n 元插值函数:

$$z(x_1, x_2, \dots, x_n) = \sum_{j=1}^m \prod_{i=1}^n A_{ij}(x_i) z_j. \quad (1)$$

所谓的变论域方法是某些论域可以随着输入变量与输出变量的变化而改变大小.变化后的论域可以表示为:

$$X_i = [-\alpha(x_i)E_i, \alpha(x_i)E_i], \quad (2)$$

$$Z = [-\beta(z)U, \beta(z)U], \quad (3)$$

其中 $\alpha(x_i), \beta(z)$ 叫做论域的伸缩因子. X_i 与 Z 叫做初始论域.基于式(1)的变论域模糊控制可以表示成如下 n 元动态插值函数:

$$z(x(t+1)) = \beta(z(x(t))) \sum_{j=1}^m \prod_{i=1}^n A_{ij} \left[\frac{x_i(t)}{\alpha(x_i(t))} \right] z_j, \quad (4)$$

其中 $x(t) = x_i(t) (i=1, 2, 3, \dots, n)$.

3.2 Q学习算法原理(Q-learning algorithm)

Q学习是Watkins提出,并由Watkins与Dayan于1992年基本证明其收敛性^[6].Q学习利用状态-动作对值函数 $Q(s, a)$ 进行迭代,其最优策略使得期望折扣报酬总和最大.Q学习的值函数定义如下:

$$Q(s, a) = R(s, s', a) + \gamma \sum_{s' \in S} P(s'|s, a) \max_{a \in A} Q(s', a), \quad (5)$$

式中 $\gamma (0 < \gamma < 1)$ 为折扣因子,本文的 γ 值取0.9; $P(s'|s, a)$ 为状态 s 在控制动作 a 发生后转移到状态 s' 的概率; agent会根据当前状态 s 选择某一动作 a ,得到下一状态 s' 及报酬值 $R(s, s', a)$,然后根据其报酬值及式(6)迭代寻优最大Q值:

$$\begin{cases} Q^{k+1}(s_k, a_k) = \\ Q^k(s_k, a_k) + \alpha [R(s_k, s_{k+1}, a_k) + \\ \gamma \max_{a' \in A} Q^k(s_{k+1}, a') - Q^k(s_k, a_k)], \\ Q^{k+1}(\tilde{s}, \tilde{a}) = Q^k(\tilde{s}, \tilde{a}), \forall (\tilde{s}, \tilde{a}) \neq (s_k, a_k), \end{cases} \quad (6)$$

式中: Q^k 代表最优值函数 Q^* 的第 k 次迭代值; $\alpha(0 < \alpha < 1)$ 是学习因子, 它控制动作更新速度, α 值越小, 算法的搜索空间越大, 算法的稳定性越好, 本文 α 值取0.1. Q学习算法在当前状态下总是选择具有最高Q值的动作, 称为贪婪策略 π^* , 如下式:

$$\pi^*(s) = \arg \max_{a \in A} Q^k(s, a). \quad (7)$$

Q算法的动作策略选择较为关键, 合适的动作策略能提高学习的收敛速度及收敛效果. 文中采用一种基于概率分布选择动作的追踪算法^[7]来构造动作选择策略. 在该策略下, 初始状态各动作被选择的概率相等, 但随着动作值函数的不断迭代, 越高的Q值的动作被选择的概率越大, 故Q算法最终将收敛于Q矩阵代表的最优策略. 该策略概率迭代公式如下:

$$\begin{cases} P_s^{k+1}(a_g) = P_s^k(a_g) + \beta(1 - P_s^k(a_g)), \\ P_s^{k+1}(a) = P_s^k(a)(1 - \beta), \forall a \in A, a \neq a_g, \\ P_{\bar{s}}^{k+1}(a) = P_{\bar{s}}^k(a), \forall a \in A, \forall \bar{s} \in S, \bar{s} \neq s. \end{cases} \quad (8)$$

在每一状态下, 对应于Q最大值的动作称为贪婪动作, 记为 a_g . 式中 $\beta(0 < \beta < 1)$ 值的大小决定了动作搜索的速度, β 值越接近1说明控制动作策略越趋于贪婪策略, 本文 β 值取0.5. $P_s^k(a)$ 代表第 k 次迭代时状态 s 下选择动作 a 的概率. 在经过足够迭代次数的探索和利用之后, Q^k 将会以概率1收敛于最优值函数 Q^* , 最终得到一个 Q^* 矩阵表示的最优控制策略.

3.3 基于Q学习的变论域模糊控制方法(A VUFC method based on Q-learning algorithm)

利用Q学习的强搜索能力与变论域模糊控制器的优势, 设计基于Q学习的变论域模糊控制结构图如图2所示.

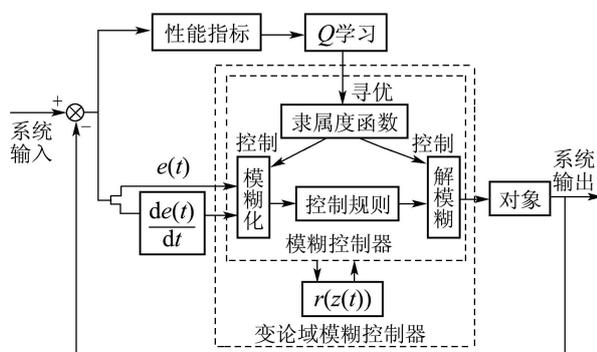


图 2 基于Q学习算法的变论域模糊控制器结构图
Fig. 2 structure of adaptive fuzzy with variable universe based on Q-learning

1) 图2中小虚线框代表模糊控制器, 输入量取误差 $e(t)$, 初始论域为 $(-E, E)$ 及误差变化率 $\frac{de(t)}{dt}$, 初始论域为 $(-EC, EC)$, 输出 z 初始论域为 $(-U, U)$.

解模糊采用重心法.

2) 大虚线框代表变论域模糊控制, 文[8]说到实现变论域控制一种常用等价方法是用在模型中量化因子 K 除以输入伸缩因子 $\alpha(x_i)$, 比例因子 K_u 乘以输出伸缩因子 $\beta(x_i)$, 这也是本文设计变论域模糊控制器采取的方法. 而在模糊规则设计好的情况下, 对变论域模糊控制要求的关键部分在于输入、输出论域伸缩因子选择, 输入、输出论域伸缩因子本文选择^[5]

$$\alpha(x, y) = \left(\frac{x}{E}\right)^{\tau_1} \left(\frac{y}{EC}\right)^{\tau_2}, \quad (9)$$

其中 $\tau_1, \tau_2 \in (0, 1)$ ^[8], 式(9)中 y 与 x 通常有着某种联系, 因此式(9)通常写为 $\alpha(x)$ 或 $\alpha(y)$, 本文取式(9)作为伸缩因子. 输出论域因子有

$$\beta(t) = \theta_i \sum_{i=1}^n P_i \int_0^t x(\tau) d\tau + \beta(0), \quad (10)$$

式(10)中 n 为输入变量个数, $\beta(0)$ 根据实际情况可以调整大小, 一般取1. θ_i, P_i 为可调参数.

在变论域模糊控制器引入等比因子后, 假定其语言分为7档, 其伸缩因子、等比因子与论域变化直观意义图参见图3所示. 在本文设计中, 用变论域模糊控制器中的伸缩因子调节论域大小, 主“外”调节; 等比因子

$$r = \frac{a}{b} = \frac{c}{d}.$$

由于无法改变论域大小, 主“内”调节, 内外结合, 以起到互补效果. 这样的设计后, 当在变论域过程中由于隶属度函数发生改变, 以致输出控制函数发生畸变时, 表现是外部性能指标(ITAE)变大, 等比因子就会配合调节, 迫使指标变小, 同时在调节中必须满足 $a + b + c + d = \alpha(x(t))E$, 这样可以保证等比因子在调节时, 使各点不会超出论域的范围.

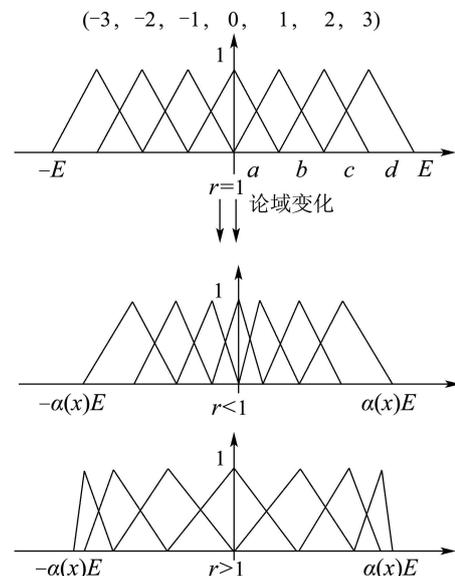


图 3 不同形式下论域变化图
Fig. 3 Different forms of variable universe

3) 性能指标本文取ITAE指标, 即

$$J(\text{ITAE}) = \int_{t_1}^{t_2} \tau |e(\tau)| d\tau = \min. \quad (11)$$

通常用这个指标的最小值来描述所设计的控制系统是否处于最优状态. 它是一个既快速平稳, 又有很好鲁棒性的性能指标. t_1 代表积分下限, t_2 代表积分上限.

4) 在本文设计的控制器的每个隶属度函数每一次动作(即每一次论域调整量)由两部分组成:

$$\Delta E = \Delta\alpha(x)E \times \Delta r, \quad (12)$$

$\Delta\alpha(x)E$ 是由变论域模糊控制器控制调整值, Δr 则是通过Q学习控制器调整. 从式(12)可以看出, 通过Q学习的调整后, 总论域调整值 ΔE 与变论域模糊控制器论域调整值 $\Delta\alpha(x)E$ 有以下关系:

$$\begin{cases} \Delta E > \Delta\alpha(x)E, \Delta r > 1, \\ \Delta E = \Delta\alpha(x)E, \Delta r = 1, \\ \Delta E < \Delta\alpha(x)E, \Delta r < 1. \end{cases} \quad (13)$$

Q学习主要是根据外部性能指标ITAE的变化来选择动作调整等比因子, 找出使指标ITAE最小的隶属度函数参数. 由此可以以ITAE值作为Q学习算法输入的状态空间集 s , 当ITAE值较大时, 即误差 e 较大时, 为使误差 e 快速进入稳定区域, 论域应进行一定的伸缩, 等比因子应该配合伸缩因子偏“外”调节, 做到粗调快速, 即 $\Delta r \geq 1$; 当ITAE值较小, 误差 e 接近稳定区域, 为了提高控制精度, 减少模糊集峰点之间的距离, 论域则应偏“内”调节, 即 $\Delta r < 1$. 每一次调整后, 观察ITAE值的变化, 同时根据奖励函数 R_i 对每次调整做出相应的评价.

Q学习控制器动作集 A 即为离散化的等比因子 Δr 值, 在算法开始阶段令等比因子 $r_0 = 1$, r 不能小于零, 故可知 Δr 的动作值划分可以为

$$A = \{-1, -0.75, -0.5, -0.25, -0.1, 0, 0.1, 0.25, 0.5, 0.75, 1\}.$$

奖励函数由于ITAE的理想目标是零, 因此在寻优时, 奖励函数 R_i 时可以设计成ITAE的实际值与目标函数值的误差平方即:

$$R_i = -(\text{ITAE} - 0)^2, \quad (14)$$

评价奖励函数主要针对ITAE而言, ITAE越大, 得到的惩罚越大, 相反, 当ITAE越接近零时, 其回报越大.

在确定奖励函数 R_i 、输入状态空间 s 和控制动作集 A 后, 根据Q算法的迭代公式就可以进行Q学习控制器在线自学习与动态优化, 其步骤如下:

1) 对所有 (s, a) 初始化各参数, 观察当前状态 $S(0)$, 令 $k = 0$.

2) 由动作概率分布在控制集中选择动作 $a(k)$.

3) 观察下一时刻的状态 $S(k + 1)$.

4) 由式(15)得到一个奖励信号 $R(k)$.

5) 根据式(7)更新Q矩阵.

6) 按照式(8)计算贪婪动作 $a_g(k)$.

7) 根据式(9)更新动作概率分布.

8) $k = k + 1$, 反回步骤2).

4 仿真算例(Simulation)

4.1 典型二阶系统(Typical second order systems)

传递函数为 $H(s) = 1/(400s^2 + 40s + 1)$, 输入、输出语言变量分为7档, 输入取误差及误差变化率, 隶属度函数全部取三角函数, 模糊规则如表1(NB: 负大, NM: 负中, NS: 负小, ZE: 零, PS: 正小, PM: 正中, PB: 正大), 采样周期 $T = 0.1$ s. 在相同条件下, 当输入信号为周期100s的正弦函数时, 同时用变论域模糊控制算法与本文设计方法对正弦信号进行跟踪. 两者输出-时间曲线图4所示. 伸缩因子参数选择: $E = EC = 2, U = 1, \tau_1 = \tau_2 = 0.2, \theta_i = 5, P_i = 5$.

表1 模糊控制规则表

Table 1 Fuzzy control rules

E	EC						
	NB	NM	NS	ZE	PS	PM	PB
NB	PB	PB	PB	PB	PM	ZE	ZE
NM	PB	PB	PB	PB	PM	ZE	ZE
NS	PM	PM	PM	PM	ZE	NS	NS
ZE	PM	PM	PS	ZE	NS	NM	NM
PS	PS	PS	ZE	NM	NM	NM	NM
PM	ZE	ZE	NM	NB	NB	NB	NB
PB	ZE	ZE	NM	NB	NB	NB	NB

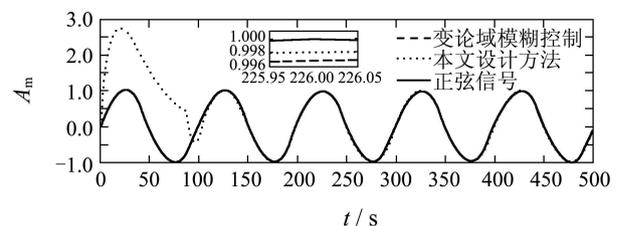


图4 二阶系统正弦信号响应过程

Fig. 4 The sinusoid responds of the two order linear process

从图4看来, 由于开始阶段是Q算法寻优阶段, 跟踪效果较变论域模糊控制器来说是非常差的, 其ITAE指标(每10s计算一次)也是非常大的, 最大达1300, 当收敛后即寻优结束后, 控制效果就会比变论域模糊控制器好些了, 其ITAE的值也降到了2.5.

为了对比两者在稳态上的差别, 也可通过ITAE指标来反映, 由于本文的设计正弦函数周期为100s, 因此可取某个时段波形, 计算时间与误差绝对值

在100 s, 因此可取某个时段波形, 计算时间与误差绝对值在100 s内的积分如200~300 s, 结果本文设计方法ITAE为100, 变论域模糊控制器为190, 图4中的中间小图为一个峰值时刻的局部放大图。

4.2 典型非最小相位系统(Typical nonminimum phase system)

水轮机调速系统属非线性、高阶、时变系统, 而且存在典型的非最小相位环节, 相对来说, 系统不易稳定且结构复杂^[9,10]. 它对控制策略的要求是: 1) 必须能实现较高精度; 2) 出现扰动时能尽快随着扰动变化实现自适应调整; 3) 该控制策略要具备较强的鲁棒性与实现性。

对于理想的无损耗的水轮机而言, 输入水轮机机械功率 ΔP_m 与导叶位置的变化 $\Delta\mu$ 满足:

$$\frac{\Delta P_m}{\Delta\mu} = \frac{1 - T_w s}{1 + 0.5 T_w s} \quad (15)$$

这就是水轮机的传递函数, 调整器辅助伺服传递函数为 $1/(T_G s + 1)$, 发电机负荷部分传递函数 $1/(T_m s + K_D)$, 因此可以得到水轮机调速器的一般模型如图5。

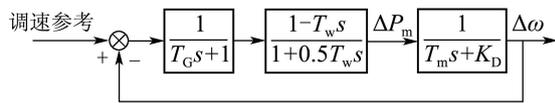


图 5 水轮机调速系统结构模型图

Fig. 5 The structure of governing system of hydro-generator

图5中: T_w 为水流时间常数, 通常取值在0.5~4 s, K_D 为系统阻尼常数, T_G 为发电机伺服时间常数, T_m 为机组惯性常数. 同二阶系统设计一样, 取两个输入分别为误差及误差变化率, 隶属度函数为7个三角函数, 模糊规则表可参考表1. 根据文[11]取

$T_w = 1.6$ s, $T_G = 0.05$ s, $T_m = 8.5$ s, $K_D = 1.3$ s, 伸缩因子参数选择为:

$$E = EC = 3, U = 3, \\ \tau_1 = \tau_2 = 0.15, \theta_i = 2, p_i = 5.$$

在无负载干扰的情况下施以2.5 Hz(5%)的频率扰动, 以作为阶跃输入的动态响应, 仿真结果如图6所示($T_w = 1.60$ s, $T_m = 8.50$ s, $K_D = 1.30$). 水轮机的一项重要内容是根据系统被控制对象的参数而调节控制器参数, 它们的不精确会对水轮机调节带来重要影响. 同时为了对比这两种算法的自适应能力, 通过改变其中某一个参数来做对比, 具体到水轮机算例上, 由于水流时间常数 T_w 对水轮机控制系统动态性能有较大的影响, 而机组习惯性常数 T_m 、阻尼系数 K_D 影响系统的稳定性, 故选取下面3组不同工

况数值进行了研究, 见表2. 其仿真结果分别对应于图7($T_w = 3.00$ s, $T_m = 8.50$ s, $K_D = 1.30$)、图8($T_w = 1.60$ s, $T_m = 11.00$ s, $K_D = 0.76$)和图9($T_w = 3.00$ s, $T_m = 11.00$ s, $K_D = 0.76$).

表 2 不同工况下传递系数表

Table 2 Transfer coefficient in different conditions

工况	传递系数		
	T_w/s	T_m/s	K_D
工况①	1.60	8.50	1.30
工况②	3.00	8.50	1.30
工况③	1.60	11.00	0.76
工况④	3.00	11.00	0.76

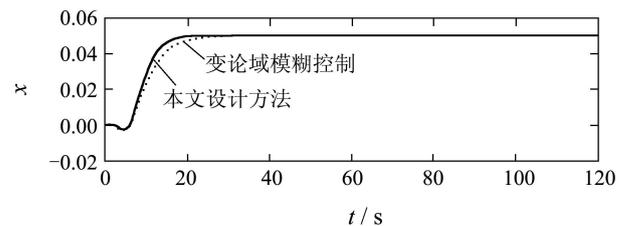


图 6 工况①下2.5 Hz频率扰动阶跃响应

Fig. 6 Step response curves of 2.5 Hz frequency for ①

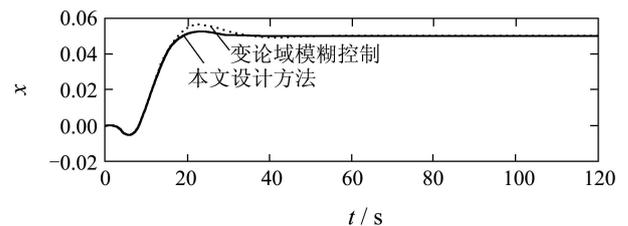


图 7 工况②下2.5 Hz频率扰动阶跃响应

Fig. 7 Step response curves of 2.5 Hz frequency for ②

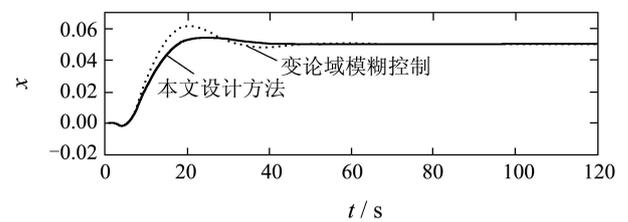


图 8 工况③下2.5 Hz频率扰动阶跃响应

Fig. 8 Step response curves of 2.5 Hz frequency for ③

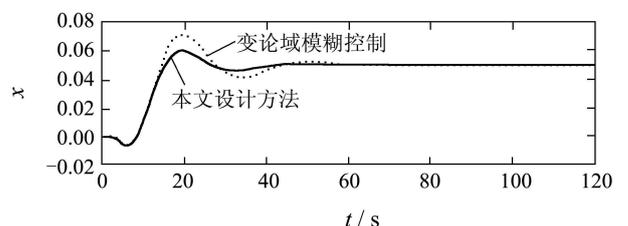


图 9 工况④下2.5 Hz频率扰动阶跃响应

Fig. 9 Step response curves of 2.5 Hz frequency for ④

从图6来看,在2.5 Hz扰动下,两个调好的控制器控制性能相差不大;但是,当参数发生变化后,从图7~图9显示来看,变论域模糊控制器的自适应能力相对本文设计算法来说就要差很多了,这一点从表3的调节时间数值以及超调量值就可以体现出来.表3中,A表示变论域模糊控制,B表示本文方法.而上述表中的ITAE数据虽是前100s的计算值,但是可以说明的是:接近稳态的时间越短其值肯定越小,从其数据来,本文设计方法也优越于变论域模糊控制.

表3 性能指标对照表
Table 3 Performance index

性能指标	工况①		工况②		工况③		工况④	
	A	B	A	B	A	B	A	B
调节时间/s	19	15	36	24	48	35	58	42
超调量/%	0	0	13.2	5.4	22	9.6	39.4	12
ITAE	4	2.6	5.4	3.3	7	5.8	11.2	7.5

5 结论(Conclusion)

变论域模糊控制器是一种高精度模糊控制器,可其控制函数在“遗传复制”中存在“失真”问题,而问题的实质是在于隶属度函数的变化引起了输入输出的变化,因此,为了减少隶属度函数的变形而使控制器的输出函数发生畸变的概率,本文提出了基于Q学习算法的变论域模糊控制算法,通过对典型二阶系统与典型非最小相位系统以及水轮机传递系数变化的基础上进行建模分析,可以得出本文算法有以下优点:

1) 它继承了变论域模糊控制器应用环境的优点,即不需要精确的数学模型、复杂时变、非线性系统,同时该算法在寻优过程中具有较快的收敛速度.

2) 本算法中,等比因子的提出分担了变论域模糊控制器中伸缩因子的“压力”,同时发现,当变论域模糊控制器中伸缩因子不够理想的情况下,根据ITAE指标,通过Q学习算法对等比因子的寻优,一样可以找到适合指标下的一组因子参数.

3) 本文算法也可以看作是对变论域模糊控制算法的一种提升,即用等比因子配合伸缩因子调节论域,极大程度上克服了由于变论域模糊控制器“失真”问题所带来对控制器动态性能的影响.在稳态下,可以说,本文设计方法在控制精度上精确度是大于或等于变论域模糊控制的.

参考文献(References):

[1] 李洪兴. Fuzzy控制的本质与一类高精度Fuzzy控制器的设计[J].

控制理论与应用, 1997, 14(6): 868 - 876.

(LI Hongxing. The essence of fuzzy control and a kind of fine fuzzy controller[J]. *Control Theory & Applications*, 1997, 14(6): 868 - 876.)

[2] 李洪兴. 变论域自适应模糊控制器[J]. 中国科学(E辑), 1999, 29(1): 32 - 42.

(LI Hongxing. Variable universe adaptive fuzzy controller[J]. *Science China(Series E)*, 1999, 29(1): 32 - 42.)

[3] 李洪兴. 非线性系统的变论域稳定自适应模糊控制[J]. 中国科学(E辑), 2002, 32(2): 211 - 223.

(LI Hongxing. Variable universe stable adaptive fuzzy control of non-linear system[J]. *Science in China(Series E)*, 2002, 32(2): 211 - 223.)

[4] 岳士弘, 张绍杰, 李平. 变论域自适应模糊控制器失真率的计算[J]. 控制理论与应用, 2005, 22(5): 807 - 809.

(YUE Shihong, ZHANG Shaojie, LI Ping. Computation of distortion ratio of adaptive fuzzy controller based on variable universe[J]. *Control Theory & Applications*, 2005, 22(5): 807 - 809.)

[5] 李家炜. 一种新的优化模糊控制器算法[J]. 控制理论与应用, 2002, 19(2): 279 - 281.

(LI Jiawei. A new optimization method for fuzzy controller's design[J]. *Control Theory & Applications*, 2002, 19(2): 279 - 281.)

[6] WATKINS J C H, DAYAN P. Q-learning[J]. *Machine Learning*, 1992, 8(3): 279 - 292.

[7] SUTTON R S, BARTO A G. *Reinforcement Learning: an Introduction*[M]. Cambridge: MIT Press, 1998: 87 - 160.

[8] 郑宏, 徐红兵, 朱贵平. 变论域自适应模糊控制在航机发电中的应用[J]. 控制理论与应用, 2008, 25(2): 253 - 256.

(ZHENG Hong, XU Hongbing, ZHU Guiping. Adaptive fuzzy control based on variable universe and its application to aero engine turbine power generator[J]. *Control Theory & Applications*, 2008, 25(2): 253 - 256.)

[9] 刘志坚, 束洪春, 于继来, 等. 一种满意控制的水轮机调速系统参数优化方法[J]. 中国电机工程学报, 2009, 29(20): 99 - 105.

(LIU Zhijian, SHU Hongchun, YU Jilai, et al. A method to optimize parameters of the hydraulic turbine governor with satisfactory control[J]. *Proceedings of the CSEE*, 2009, 29(20): 99 - 105.)

[10] 景雷, 叶鲁卿. 面向目标的水轮发电机组智能模糊控制[J]. 中国电机工程学报, 1996, 16(1): 63 - 67.

(JING Lei, YE Luqing. New intelligent fuzzy control system for hydroelectric generating units[J]. *Proceedings of the CSEE*, 1996, 16(1): 63 - 67.)

[11] 寇攀高, 周建中, 何耀耀, 等. 基于菌群-粒子群算法的水轮发电机组PID调速器参数优化[J]. 中国电机工程学报, 2009, 29(26): 101 - 106.

(KOU Pangao, ZHOU Jianzhong, HE Yaoyao, et al. Optimal PID governor tuning of hydraulic turbine generators with bacterial foraging particle swarm optimization algorithm[J]. *Proceedings of the CSEE*, 2009, 29(26): 101 - 106.)

作者简介:

余涛 (1974—), 男, 副教授, 博士, 主要研究领域为复杂电力系统的非线性控制理论和仿真等, E-mail: taoyu1@scut.edu.cn;

于文俊 (1986—), 男, 硕士研究生, 主要研究方向为电力系统优化控制, E-mail: yuwenjun1002@126.com;

李章文 (1988—), 男, 硕士研究生, 主要研究方向为电力系统优化控制, E-mail: lizhangwen@qq.com.