Информационные процессы, Том 9, № 1, 2009, стр. 44–50.

Теоретические основы построения многопараметрических функций принадлежности нечетких систем

Е.А. Халов

Липецкий государственный технический университет, Липецк, Россия
Поступила в редколлегию 24.03.2009

Аннотация—Излагается способ расширения класса кривых, пригодных для характеризации многопараметрических функций принадлежности нечетких систем. Анализируется возможность использования алгебраических многочленов определенного вида для построения функций принадлежности. Рассматривается доказательство пригодности таких многочленов для построения функций, наименее уклоняющихся от экспертных функций принадлежности.

1. ВВЕДЕНИЕ

Функции принадлежности (ФП) $X(x)$ являются единственным средством описания нечетких подмножеств [1]. При этом численное значение степени принадлежности, обозначаемое в виде $X(x)$, характеризует степень принадлежности элемента x некоторому нечеткому подмножеству X, являющемуся в выражении естественного языка некоторой характеристикой технологического процесса (ТП) — температурой металла, скоростью движения раската, давлением воды и т.д. — и позволяет адекватно отразить мнение экспертов в связи с неспособностью человека формулировать свою количественную оценку в виде одиночного (точного) числа.

В обзоре [2] хорошо представлена “эволюция” аналитических выражений ФП, имеющая одну примечательную особенность: первоначально системы нечеткого вывода использовали лишь линейные ФП (γ–, t–класса и линейные трапецеидальные), а затем, исчерпав их возможности, начались исследования и попытки применения нелинейных ФП — вначале со сравнительно простыми аналитическими выражениями (сигмоидальные и гауссовы семейства функций), а в дальнейшем и сложных нелинейных зависимостей (z–, s–, π–образных и нелинейных трапецеидальных) в стремлении улучшить их характеристики (гладкость, количество управляющих параметров и др.).

В то же время изучение ФП на протяжении истории развития нечетких систем, наблюдаемой по доступным для изучения публикациям, носит весьма пассивный характер: совершенствование процесса функционирования нечетких систем в большинстве случаев вообще не затрагивало этап фаззификации или совершенствовало его весьма незначительно. В итоге, ФП оставалась (и остается) малоисследованным компонентом, несмотря на важную роль этапа фаззификации при функционировании нечетких систем.

Следует отметить, что в основополагающей работе Л. Заде [3], опубликованной в 1965 году, были введены два важнейших понятия:

- функции принадлежности;
- многозначных операторов, предназначенных для работы с ФП,

определивших дальнейшее развитие теории нечетких множеств. Но к сожалению, среди множества публикаций, вышедших после появления этой работы, подавляющее большинство исследований было посвящено лишь различным обобщениям операторов, предложенных Заде в [3] и лишь немногие исследователи уделяли внимание изучению собственно функций принадлежности. В настоящее время, на основе анализа доступных публикаций в отечественных и зарубежных сборниках трудов, а также в специализированных журналах (“Fuzzy Sets and Systems”, “Information Sciences”, “Information and Control”, серии периодических журналов издательства “IEEE”, “Springer” и подобных им, вплоть до последних вышедших номеров), к сожалению, выявляется аналогичная тенденция. Более того, в большинстве работ функции принадлежности как таковые вообще не упоминаются. Но тем не менее, нельзя забывать о факте происхождении понятия функции принадлежности [3, 4] и его исключительной важности в теории нечетких подмножеств, а также о том, что именно аппарат функций принадлежности позволяет характеризовать лингвистические термы в базе нечетких правил. В свою очередь, база нечетких правил может трактоваться как некоторое разбиение пространства влияющих факторов на области с размытыми (нечеткими) границами, в каждой из которых функция отклика принимает значение, заданное соответствующей функцией принадлежности, а отдельное правило в базе знаний представляет собой “информационный сгусток”, отражающий одну из особенностей зависимости “входы–выходы”. Как отмечал Заде, такие “сгустки насыщенной информации” или “гранулы знаний” могут рассматриваться как аналог вербального кодирования или накопления опыта, которое происходит в человеческом мозге при обучении. С этой точки зрения, недостаточное внимание, уделяемое функциям принадлежности, сравнимо, например, с игнорированием функции распределения вероятности в теории вероятностей. Наконец, процесс построения ФП на сегодняшний день не формализован, а в качестве аналитических выражений, характеризующих функции принадлежности, спонтанно используются лишь широко известные математические функции или функционалы [2].

В этой связи представляется целесообразной выработка принципиально иного подхода к построению ФП, не связанного с использованием одного (реже двух) фиксированных аналитических выражений в их описании. Следует также отметить, что настоящая работа является дальнейшим развитием подхода, предложенного в [8].

2. ПОСТАНОВКА ЗАДАЧИ

В качестве объекта исследования рассмотрим структурно–параметрическую модель функции принадлежности, максимально достоверно отражающую высказывания (предпочтения) человека–эксперта в процессе принятия им решений в контуре “человек–технологический процесс”. С точки зрения автора данной статьи, ключевая особенность такого подхода к формулированию постановки задачи состоит в том, что к рассмотрению принимается **модель** функции принадлежности, которая будет являться приближением “настоящей” (эталонной) функции принадлежности, формируемой экспертом. Отметим, что иногда такое приближение оказывается весьма точным.

С одной стороны, совершенно очевидно, что модель такой ФП осуществляет преобразование входного значения технологического параметра x в выходное значение $X(x)$ при помощи некоторого оператора f (компоненты которого впоследствии подвергаются идентификации), выражаемое зависимостью вида:

$$X(x) = f(x), \tag{1}$$

где

x — входное вещественное значение технологического параметра;

$X(x)$ — выходное вещественное значение степени принадлежности, вычисленное при помощи модели ФП.

С другой стороны, реальные высказывания, формируемые экспертом с учетом накопленного положительного опыта работы с ТП (и, что важно — в результате неформализуемого мыслительного процесса и собственной интуиции), также подчиняются некоторой априори неизвестной зависимости, характеризуемой, в свою очередь, "одушевленным" оператором f^*:

$$X^*(x) = f^*(x), \tag{2}$$

где $X^*(x)$ — вещественное значение степени принадлежности, формируемое человеком–экспертом.

Здесь, под "одушевленным" оператором f^* будем понимать весьма сложное и слабоформализуемое правило, присущее исключительно человеку (эксперту, инженеру — т.е. лицу, принимающему решение), согласно которому *сигналы*, поступающие с датчиков ТП (относящиеся к конкретному технологическому параметру), преобразуемые при помощи измерительного оборудования в *данные* и воспринимаемые органами чувств оператора уже в форме осмысленной *информации*, в итоге трансформируются (экспертом) в неточные и приблизительные (нечеткие) значения степеней принадлежности. Важно отметить, что в рамках данной работы принимается допущение, состоящее в том, что *неточные и приблизительные значения степеней принадлежности*, формируемые экспертом в связи с его неспособностью сформулировать свою количественную оценку в виде одиночного числа [1], заменяются *точным значением степени принадлежности*. Такое допущение, а по сути — огрубление значений оценок, выдаваемых экспертом, на данном этапе исследований позволит избежать усложнений при вычислении значений степеней принадлежности, связанных с использованием интервальной арифметики (которая, тем не менее, представляется более адекватной для восприятия человеком).

Далее, кроме входной переменной x, результат вычисления по модели ФП определяется также выбором алгоритма $\Psi_{X(x)}$ вычисления значений степеней принадлежности $X(x)$, компонентов структуры $\mathbb{Q}$ модели и вектора управляющих параметров $\mathbb{C}$, составляющих оператор f модели функции принадлежности:

$$f = \langle \Psi_{X(x)}, \mathbb{Q}, \mathbb{C} \rangle. \tag{3}$$

Задачей настоящего исследования будет являться построение доказательства существования такой функции принадлежности $X(x)$, для которой оператор f станет максимально точно соответствовать оператору f^*, характеризующему высказывания человека–эксперта. Иными словами, должно соблюдаться условие близости (в некотором смысле) рассматриваемых операторов:

$$f \approx f^*. \tag{4}$$

Количественной характеристикой степени близости операторов f и f^* может служить величина невязки Δ между величиной степени принадлежности $X^*(x)$, формируемой экспертом и выходной величиной $X(x)$, вычисленной при помощи модели функции принадлежности

$$\Delta = X^*(x) - X(x) \tag{5}$$

при одинаковом значении величины технологического параметра x.

Согласно результатам проведенных исследований можно сделать предварительный вывод о существовании искомой функции принадлежности $X(x)$, для которой выполняется условие (4). Забегая немного вперед, отметим, что предварительное построение и визуализация (при помощи MS Excel 2003) заданного графика функции принадлежности $X^*(x)$ и подбор необходимого многочлена и его коэффициентов на интервале носителя $\mathbf{X}_0$, позволяет убедиться в верности выдвинутого предположения: “зазор” между графиками функций принадлежности $X^*(x)$ и $X(x)$ уменьшается. В дальнейшем, в целях упрощения, функцию принадлежности $X^*(x)$ будем именовать экспертной функцией принадлежности.

Этот экспериментальный факт дает основание полагать, что последовательность многочленов определенного вида, характеризующих непрерывную экспертную функцию принадлежности $X^*(x)$, равномерно сходится к ней.

Расширяя классы допустимых кривых, пригодных для характеризации функции принадлежности $X(x)$, описание предлагается проводить в форме обобщенного многочлена, имеющего следующую структуру:

$$\Phi^n(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots + a_i\varphi_i(x) + \ldots + a_{n-1}\varphi_{n-1}(x) + a_n\varphi_n(x), \tag{6}$$

где

$\Phi^n(x)$ — обобщенный многочлен степени (порядка) n;

$\varphi_i(x)$ — базисная функция, $i = \overline{1,\ n}$.

Заметим, что для описания также возможно использовать и некоторые нелинейные комбинации функций, отличные от многочлена в выражении (6).

Если в качестве базисных функций использовать степенные функции вида $\varphi_i(x) = x^i$, то возникает задача приближения алгебраическими многочленами (полиномами) n-ной степени [9]. При этом наиболее часто используется стандартная степенная форма представления таких многочленов, именуемая мономиальным базисом:

$$\Phi^n(x) = a_0 + a_1x + \ldots + a_ix^i + \ldots + a_{n-1}x^{n-1} + a_nx^n = \sum_{i=0}^{n} a_ix^i, \tag{7}$$

где

a_i — фиксированный коэффициент многочлена, $i = \overline{0,\ n}$, $a_i \in \mathbf{R}$;

a_n — старший коэффициент, $a_n \neq 0$.

Применение таких многочленов в большинстве случаев приводит к нежелательным осцилляциям значений $\Phi^n(x)$. Причина подобного поведения интерполяционного многочлена заключается в том, что он по сути представляет собой сумму степенных функций. Эти функции обладают особым свойством — их значения на всем интервале $\mathbf{X}_0$ определяются значением на произвольном малом отрезке этого интервала. Подбирая коэффициенты a_i, $i = \overline{0,\ n}$ многочлена $\Phi^n(x)$ таким образом, чтобы сумма составляющих его функций a_ix^i принимала в нескольких точках искомые значения, становится невозможным контролировать значения отдельных членов в остальных точках $i = \overline{0,\ n}$. Поскольку значение каждого слагаемого может быть достаточно велико, то возможны значительные отклонения значений принадлежности $X(x)$, выходящие за пределы интервала значений степени принадлежности $[0,\ 1]$. Также следует отметить, что многочлены степеней $n > 5$ редко применяются при построении численных алгоритмов. К тому же выражение (7) содержит значительное количество коэффициентов a_i, $i = \overline{0,\ n}$, физический и геометрический смысл которых подчас трудно осмыслить, следовательно, для конструирования сколь бы то ни было сложных ФП с хорошей управляемостью они не пригодны.

С точки зрения простоты и вычислительной эффективности целесообразно использовать многочлены невысоких степеней ($n = 2$, $n = 3$). Как уже было сказано выше, привлекая в качестве $\Phi^n(x)$ многочлены больших степеней, можно генерировать весьма сложные зависимости, а следовательно, и весьма сложные функции принадлежности. Но, как уже было сказано выше, такие многочлены содержат значительное количество коэффициентов, имеющих малопонятный смысл, и, что наиболее важно, возникают сопутствующие им нежелательные осцилляции значений степени принадлежности, построенной на их основе. Следовательно, необходимо отыскать такой способ построения модели функции принадлежности $X(x)$, который позволит это сделать с *наименьшим уклонением* от заданной эталонной (экспертной) функции принадлежности $X^*(x)$.

В последнее время во многих инженерных дисциплинах растет интерес к специфическим, но менее осциллирующим аналитическим функциям для решения широкого круга практических задач. Расширяя таким образом допустимые классы кривых, пригодных для характеризации ФП, сосредоточим внимание на особом классе многочленов, именуемых *многочленами Бернштейна*, которые представляют собой алгебраические многочлены в виде линейной комбинации базисных многочленов Бернштейна. Математические аспекты использования таких многочленов в математических исследованиях были описаны Сергеем Натановичем Бернштейном еще в 1912 году в конструктивном доказательстве аппроксимационной теоремы Вейерштрасса [7]. Кроме того, они достаточно подробно представлены в других источниках [9, 10, 11, 12]. Теперь перейдем к более подробному рассмотрению многочленов Бернштейна.

На интервале [0, 1] рассмотрим фиксированный набор из $(n+1)$ точек, именуемых *узловыми*. Значение координаты каждой такой точки x_k будет определяться в соответствии со следующим соотношением: $x_k = {}^k\!/_n$, $k = \overline{0,\ n}$.

Согласно Бернштейну [7, 12], многочлен (n-й степени) в операторе f модели функции принадлежности $X(x)$, описывающей на этом интервале экспертную функцию принадлежности $X^*(x)$, будет описываться скалярным выражением, имеющим следующую структуру:

$$\Phi^n(x) = \sum_{k=0}^{n} X^*\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}.$$

где $C_n^k x^k (1-x)^{n-k}$ — базисный многочлен (базис) Бернштейна, $k = \overline{0,\ n}$.

Для многочленов Бернштейна в [6] получены некоторые полезные соотношения, которые будут необходимы при изложении доказательства:

$$\begin{aligned} &\sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} = 1; \quad \sum_{k=0}^{n} k C_n^k x^k (1-x)^{n-k} = nx; \\ &\sum_{k=0}^{n} (k-nx)^2 C_n^k x^k (1-x)^{n-k} = nx(1-x) \le \frac{n}{4}. \end{aligned} \tag{8}$$

Таким образом, при $X(x) = \Phi^n(x)$ нам предстоит доказать, что для заданного отклонения $\varepsilon > 0$ существует такое n, что

$$|X^*(x) - X(x)| < \varepsilon, \quad \forall x \in [0,\ 1]. \tag{9}$$

Теперь перейдем к изложению доказательства неравенства (9).

Доказательство существования искомой ФП изложим, следуя Бернштейну [6] и опираясь на теорему Вейерштрасса о приближении [5]. Согласно данной теореме можно утверждать, что любая непрерывная на интервале носителя $\mathbf{X}_0$ функция принадлежности $X^*(x)$ описывается равномерно сходящейся на этом интервале последовательностью многочленов. Обратное утверждение также верно, поэтому теорема Вейерштрасса устанавливает характеристическое свойство функции принадлежности как непрерывной функции. Это подтверждает представление о функции принадлежности $X^*(x)$, формируемой экспертом, как о некотором аналитическом выражении и дает основание полагать, что функция принадлежности $X^*(x)$, будучи непрерывной, близка к классу многочленов. Также, данное обстоятельство оправдывает интуитивное предположение о близости операторов f и f^* в выражении (4), а следовательно, окончательно обосновывает целесообразность использования многочленов в структуре оператора f при построении модели ФП.

Для удобства изложения доказательства, вместо произвольного (обычно положительного) интервала вещественной оси, заключающего в себе носитель $\mathbf{X}_0$ функции принадлежности, используем единичный интервал $[0,\ 1]$. В дальнейшем, при помощи соответствующей операции масштабирования (нормировки), можно будет осуществить обратный переход от единичного интервала к произвольному интервалу носителя $\mathbf{X}_0$.

Пусть на указанном интервале каким-либо способом определена экспертная функция принадлежности $X^*(x)$. Обозначим через $X^*_{\max}(x)$ наибольшее значение степени принадлежности экспертной функции принадлежности $X^*(x)$ на указанном единичном интервале $[0,\ 1]$ носителя $\mathbf{X}_0$.

Согласно условию равномерной непрерывности, сформулированному в теореме Кантора, если функция $X(x)$ непрерывна на отрезке $[a,\ b]$, то для любого $\varepsilon > 0$ существует $\delta > 0$ со свойством:

$$\left|x - x'\right| < \delta \quad \left(x,\ x' \in \mathbf{X}_0\right) \quad \Rightarrow \quad \left|X(x) - X\left(x'\right)\right| < \varepsilon. \tag{10}$$

Подберем $\delta > 0$ так, чтобы условие непрерывности (10) теперь выполнялось при значении $\varepsilon/2$ вместо ε, а также выберем произвольное значение n, удовлетворяющее условию $n > \frac{X^*_{\max}(x)}{\varepsilon\,\delta^2}$. Тогда, согласно первому из неравенств в (8), после подстановки получим выражение следующего вида:

$$X(x) - X^*(x) = \sum_{k=0}^{n}\left[X^*\left(\frac{k}{n}\right) - X^*(x)\right]C_n^k x^k (1-x)^{n-k}.$$

Теперь зафиксируем переменную x. При этом предыдущее выражение преобразуется в неравенство вида:

$$|X(x) - X^*(x)| \le \sum_{k=0}^{n}\left[X^*\left(\frac{k}{n}\right) - X^*(x)\right]C_n^k x^k (1-x)^{n-k} = S_1 + S_2,$$

где

S_1 – результат суммирования по тем узловым точкам k, для которых выполняется условие $\left|\frac{k}{n} - x\right| < \delta$;

S_2 – результат суммирования по остальным точкам k, для которых $\left|\frac{k}{n} - x\right| \ge \delta$, что также эквивалентно неравенству $1 \le \frac{(k-n\,x)^2}{(n\,\delta)^2}$.

Из выбора δ и условия непрерывности (10) получаем, что $S_1 < \frac{\varepsilon}{2}$. Далее,

$$1 \le \frac{(k-nx)^2}{(n\delta)^2} \quad \Rightarrow \quad S_2 \le \frac{2X^*_{\max}(x)}{(n\delta)^2}\sum_{k=0}^{n}(k-nx)^2 C_n^k x^k (1-x)^{n-k}. \tag{11}$$

Исходя из соотношений (8), (11) и выбора степени n, получаем, что $S_2 < \frac{\varepsilon}{2}$. Следовательно, неравенство (9) доказано.

Таким образом, можно считать обоснованным предположение о возможности использования многочленов $\Phi^n(x)$, обобщенный вид которых приведен в выражении (6), в структуре оператора f, который составляет модель функции принадлежности из выражения (3). В частности, одними из представителей данного класса многочленов являются многочлены Бернштейна, обладающие несомненными положительными свойствами, проявление которых можно удачно сочетать с некоторыми требованиями, накладываемыми на функции принадлежности [13].

3. ЗАКЛЮЧЕНИЕ

Представленный материал подтверждает, что выбор многочленов Бернштейна для целей построения функций принадлежности неслучаен и вполне обоснован. Так, основными положительными особенностями этих многочленов являются их вычислительная устойчивость [10] и неотрицательность на всем протяжении носителя $\mathbf{X}_0$. Как показали проведенные вычислительные эксперименты, оставшиеся за рамками настоящей статьи, использование многочленов Бернштейна позволяет формировать функции принадлежности, обладающие широкими интерактивными возможностями настройки параметров своего положения и формы.

На сегодня есть основания полагать, что подход к построению функций принадлежности, предложенный автором в [8] и получивший теоретическое обоснование в настоящей работе, будет развиваться и далее, а также окажется перспективным и востребованным при построении нечетких систем моделирования и управления технологическими процессами.

СПИСОК ЛИТЕРАТУРЫ

1. Заде Л.А. Тени нечетких множеств. *Проблемы передачи информации*, 1966, том 2, вып. 1, стр. 37–44.
2. Халов Е.А. Одномерные многопараметрические функции принадлежности в задачах нечеткого моделирования и управления. *Мехатроника, автоматизация, управление. Приложение*, 2007, № 4, стр. 2–11.
3. Zadeh L.A. Fuzzy Sets. *Information and Control*, 1965, vol. 8, issue 3, pp. 338–353.
4. Заде Л.А. *Понятие лингвистической переменной и его применение к принятию приближенных решений.* М.: Мир, 1976.
5. Корн Г. *Справочник по математике для научных работников и инженеров.* М.: Наука, 1974.
6. Натансон И.П. *Теория функций вещественной переменной.* М.: Наука, 1974.
7. Бернштейн С.Н. *Собрание сочинений* в 3 т. М.: Изд–во АН СССР, 1952. Том 1: *Конструктивная теория функций.* 1954.
8. Халов Е.А., Гвозденко Н.П. О расширении класса допустимых кривых для построения функций принадлежности. *Управление большими системами: сборник трудов V Всероссийской школы-семинара молодых ученых.* Липецк: ЛГТУ, 2008, том 1, стр. 92–99.
9. Волков Е.А. *Численные методы: учеб. пособие для вузов.* М.: Наука, 1987.
10. Farouki R.T. and Goodman T.N.T. On the Optimal Stability of the Bernstein Basis. *Mathematics of Computation.* 1996, vol. 65, no. 216, pp. 1553–1566.
11. Farouki R.T. and Rajan V.T. Algorithms for Polynomials in Bernstein Form. *Computer Aided Geometric Design*, 1988, vol. 5, issue 1, pp. 1–26.
12. Виденский В.С. *Многочлены Бернштейна.* Л.: ЛГПИ им. А.И. Герцена, 1990.
13. Халов Е.А. Виды и свойства одномерных функций принадлежности нейро-нечетких систем моделирования и управления. *Вестник ЛГТУ–ЛЭГИ*, 2008, № 1, стр. 32–43.