УДК 004.94

К. А. Майков, С. М. Жиряков

МЕТОД УНИВЕРСАЛЬНОЙ АППРОКСИМАЦИИ В НЕЧЕТКОМ ВЫВОДЕ, ИСПОЛЬЗУЮЩИЙ ЛИНЕЙНЫЕ ОБОЛОЧКИ КОРРЕКТИРУЮЩИХ ДАННЫХ

Предложена модификация метода нечеткого вывода решения, позволяющая без потери объяснительных возможностей осуществлять корректировку значения выводимой переменной в указываемых окрестностях точек с нарушениями требований к точности решения задачи.

E-mail: maikov@mx.bmstu.ru; zs-mailbox@mail.ru

Ключевые слова: *нечеткая логика, фаззификация, модификация алгоритма Суджено.*

Известно, что по сравнению с рядом известных аналогов (статистические методы, метод нейронных сетей) нечеткий вывод как метод поддержки принятия решения обладает такими преимуществами, как объяснение получаемого решения, возможность обобщения решения и работоспособность в условиях недостоверности или зашумленности в исходных данных. Наряду с этим один из основных недостатков нечеткого вывода заключается в его невысокой точности, обусловленной значительным влиянием субъективного фактора при формировании характеристических функций лингвистических переменных и правил взаимосвязи переменных между собой. Однако существуют доказательства, предложенные в [1–3], показывающие универсальную аппроксимационную способность метода нечеткой логики.

В данной статье предлагается метод, который позволяет совместить преимущества объяснения решения в терминах, вводимых экспертом предметной области, и определенных им видах характеристических функций с возможностью аппроксимации к некоторой функциональной зависимости. При этом точность аппроксимации зависит от объема корректирующих данных.

Будем считать, что эксперт на основе собственных представлений о предметной области формулирует продукционные правила поведения исследуемой системы в виде

$$\text{ЕСЛИ } (X = X_1) \text{ И } (Y = Y_1) \text{ ТО } (Z = Z_1), \tag{1}$$

где $X,\ Y,\ Z$ — лингвистические переменные; $X_1,\ Y_1,\ Z_1$ — термы.

Рассмотрим схему нечеткого вывода, общую для выводов Tsukamoto [4], Sugeno [5] и упрощенного вывода [6] на примере правила (1).

На первом этапе фаззификации определяют степень принадлежности $\alpha_x,\ \alpha_y$ четких исходных данных (x_0, y_0) к соответствующим термам X_1 и Y_1, указанным в левой части правила. На следующем этапе нечеткого вывода — логическом выводе — эти степени принадлежности $\alpha_x,\ \alpha_y$ определяют с помощью преобразования ξ четкого значения $z = \xi(\alpha_x, \alpha_y)$ выводимой переменной. Полученное значение z затем участвует в композиции со значениями z_i правил вывода со степенью истинности решения $\alpha_i = \min(\alpha_x, \alpha_y)$. Композиция, как правило, осуществляется путем усреднения в соответствии с формулой взвешенного среднего

$$z_{\text{общ}} = \frac{\sum\limits_i \alpha_i z_i}{\sum\limits_i \alpha_i}\ . \tag{2}$$

Будем считать, что каждое значение z_i, получаемое из некоторого правила, отождествляется экспертом с “правильным” или желаемым решением при попадании вектора исходных данных в окрестность определения термов левой части правила, а формулой (2) выражают компромисс между имеющимися решениями z_i при формулировании итогового решения. В этом случае проблема повышения точности метода нечеткой логики может быть сведена к поиску такого преобразования ξ этапа логического вывода, которое позволит получать промежуточные значения выводимой переменной z_i, близкие к требуемому итоговому решению.

В случае вывода Tsukamoto [4] преобразование ξ характеризуется особенностями поиска решения z уравнения $\min(\alpha_x, \alpha_y) = Z_1(z)$; закономерность взаимосвязи переменных X и Y с переменной Z определяется видом терма Z_1. Если эксперт задает вид характеристической функции со смысловой точки зрения терма Z_1, то очевидно, что закономерность будет отражена искаженной, в противном случае, если эксперт будет пытаться ориентироваться на получение правильной закономерности, смысл терма Z_1 может быть потерян.

В упрощенном варианте вывода значение z принято константой [6], т.е. преобразование $\xi(\alpha_x, \alpha_y) = c$. Такой выбор приводит к упрощению взаимосвязи переменных X и Y с Z, что отрицательно влияет на точность упрощенного нечеткого вывода.

В выводе Sugeno [5] преобразование ξ задано как линейная комбинация входных данных с фиксированными коэффициентами $z = k_x x_0 + k_y y_0$. Указанная зависимость соответствует линейной аппроксимации закона взаимосвязи переменных X и Y с Z. Недостаток такого определения преобразования ξ в том, что аппроксимирующая гиперплоскость должна проходить через начало координат. Также возникает проблема выбора коэффициентов $k_x,\ k_y$ для отражения закономерности взаимосвязи переменных.

Введем новое преобразование ξ для формирования результата z этапа логического вывода при обработке продукционного правила:

$$\xi(\alpha_x, \alpha_y) = z = \alpha_x \upsilon_x + \alpha_y \upsilon_y, \tag{3}$$

где $\alpha_x = X_1(x_0)$, $\alpha_y = Y_1(y_0)$ — значения степеней истинности для каждого терма левой части правила; υ_x, υ_y — коэффициенты влияния переменных X и Y на переменную Z соответственно.

Коэффициент α_i для получаемого значения z, который участвует в композиции согласно формуле (2), определим как

$$\alpha_i = \min(\alpha_x, \alpha_y). \tag{4}$$

Покажем, что, используя формулы (3) на этапе логического вывода и (2) на этапе композиции, можно дополнить составленную экспертом систему продукционных правил новыми правилами, которые строятся с помощью набора корректировочных данных, и добиться приближения к любой непрерывной функции, задающей взаимосвязь входных и выходной переменной.

Рассмотрим область исходных данных, на которой характеристические функции термов X_1 и Y_1 принимают ненулевые значения. Для наглядности рассуждений будем полагать, что для точек исходных данных $A(x_0, y_0)$, $B(x_1, y_1)$ и $C(x_2, y_2)$ заданы требуемые значения выводимой переменной Z, равные z_0, z_1 и z_2 соответственно.

Треугольник ABC ограничивает область исходных данных, на которой можно установить линейную оболочку векторов, построенную на векторах $\overrightarrow{AM}$ и $\overrightarrow{AN}$. Подмножество векторов линейной оболочки $span\{\overrightarrow{AM}, \overrightarrow{AN}\}$, для которых справедливо

$$\vec{x} = \lambda_1 \overrightarrow{AM} + \lambda_2 \overrightarrow{AN}; \tag{5}$$

$$\lambda_1, \lambda_2 \geqslant 0; \tag{6}$$

$$\lambda_1 + \lambda_2 \leqslant 1 \tag{7}$$

будем называть *Зоной* Ψ действия поправки с *базисом Зоны* $\{\overrightarrow{a}, \overrightarrow{b}\}$, образованным корректировочными данными в некоторых точках A, B и C, где $\vec{a} = \overrightarrow{AM}$, $\vec{b} = \overrightarrow{AM}$. Точку A будем называть *основанием Зоны* — $Base(\Psi)$; а множество точек исходных данных, попадающих в треугольник ABC, — *областью определения Зоны* — $Def(\Psi)$.

Задача построения модифицированного нечеткого вывода заключается в том, чтобы при попадании произвольной точки с исходными данными $X(x^*, y^*) \in Def(\Psi)$ получать результат (значение переменной вывода), принадлежащий Зоне, что обеспечит линейную аппроксимацию поверхности, отражающей закономерность взаимосвязи переменных X и Y.

Рассмотрим точку $X(x^*, y^*) \in Def(\Psi)$, тогда в Зоне Ψ для нее существует соответствующая точка, определяемая положением вектора $\vec{x}(x^* - A_x, y^* - A_y)$. Представим $\vec{x}$ в виде разложения по векторам базиса Зоны

$$\vec{x} = \alpha\vec{a} + \beta\vec{b}. \tag{8}$$

Представим каждый из векторов $\vec{a}$,$\vec{b}$ и $\vec{x}$ в виде суммы проекции, принадлежащей пространству $Def(\Psi)$ и его ортогонального дополнения

$$\vec{x} = \vec{x}_0 + \vec{x}_\perp; \quad \vec{a} = \vec{a}_0 + \vec{a}_\perp; \quad \vec{b} = \vec{b}_0 + \vec{b}_\perp. \tag{9}$$

Тогда выражение (8) может быть преобразовано к виду

$$\vec{x} = (\alpha\vec{a}_0 + \beta\vec{b}_0) + (\alpha\vec{a}_\perp + \beta\vec{b}_\perp). \tag{10}$$

Заметим, что слагаемое $(\alpha\vec{a}_\perp + \beta\vec{b}_\perp)$ в формуле (10) равно

$$\alpha(z_1 - z_0) + \beta(z_2 - z_0) \tag{11}$$

и фактически характеризует требуемое значение выводимой переменной как

$$z_{\text{вых}} = z_0 + \|\vec{x}_\perp\| = z_0 + \alpha(z_1 - z_0) + \beta(z_2 - z_0), \tag{12}$$

поскольку $\vec{x}_\perp$ принадлежит ортогональному дополнению пространства $Def(\Psi)$.

В то же время, слагаемое $(\alpha\vec{a}_0 + \beta\vec{b}_0)$ в (10) дает разложение вектора $\vec{x}_0$ по базису из векторов $\vec{a}_0$, $\vec{b}_0$ в пространстве $Def(\Psi)$. Значения коэффициентов αи β определяются решением системы уравнений

$$\begin{pmatrix} a_{0x} b_{0x} \\ a_{0y} b_{0y} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} x_{0x} \\ x_{0y} \end{pmatrix}. \tag{13}$$

Представим решение (13) в виде

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} g_{11} g_{12} \\ g_{21} g_{22} \end{pmatrix} \begin{pmatrix} x_{0x} \\ x_{0y} \end{pmatrix}. \tag{14}$$

Тогда значение выводимой переменной согласно (12) может быть записано после преобразований в виде

$$z_{\text{вых}} = z_0 + (g_{11}a_\perp + g_{21}b_\perp)x_{0x} + (g_{12}a_\perp + g_{22}b_\perp)x_{0y}. \tag{15}$$

Таким образом, произвольная точка $X(x^*, y^*)$, характеризующая значение исходных данных и попадающая в область определения Зоны $Def(\Psi)$, задает решение с помощью формулы (15), которое принадлежит плоскости, аппроксимирующей поверхность правильных значений выводимой переменной.

Сформулируем критерий, согласно которому можно выразить условие принадлежности точки исходных данных $X(x^*, y^*)$ области определения некоторой Зоны.

Принадлежность $X(x^*, y^*)$ к $Def(\Psi)$ определяется формулами (6) и (7), однако на практике реализация условия (6) может привести к значительным вычислительным затратам, связанным с решением системы уравнений (13), количество которых равно размерности взаимосвязи входных и выходной переменных.

Рассмотрим ситуацию вычислений с учетом только условия (7). Из (14) представим сумму коэффициентов разложения вектора $\vec{x}$ в виде

$$\alpha + \beta = (g_{11} + g_{21})x_{0x} + (g_{12} + g_{22})x_{0y}. \tag{16}$$

Для уменьшения числа точек, ошибочно относимых по условию (16) к области определения Зоны $Def(\Psi)$, добавим условия в качестве компенсации потери требования неотрицательности коэффициентов α и β из условия (6)

$$\begin{aligned} \min(A_x, B_x, C_x) \leqslant x_{0x} \leqslant \max(A_x, B_x, C_x) \\ \min(A_y, B_y, C_y) \leqslant x_{0y} \leqslant \max(A_y, B_y, C_y) \end{aligned} \tag{17}$$

Благодаря условию (17) исключаются точки, ошибочно относимые к области определения Зоны $Def(\Psi)$.

При необходимости строгого выполнения условия (6), которое исключает ошибку при определении принадлежности некоторой точки к области $Def(\Psi)$, можно сформулировать дополнительные условия вида (7), используя формулу (16) для вычисления суммы коэффициентов разложения α и β.

Определим порядок построения продукционных правил, которые дополнят исходный набор, составленный экспертом предметной области. В случае попадания точки входных данных $X(x^*, y^*)$ в область определения Зоны $Def(\Psi)$ этот набор позволяет получать значение выводимой переменной, принадлежащей Зоне Ψ.

Для учета значений корректировочных данных в точках A,B и C необходимы две группы правил:

$$\text{ЕСЛИ } (\textit{Область } X = \textit{Область } X\_\Psi) \text{ И } (\textit{Область } Y = \\ = \textit{Область } Y\_\Psi) \text{ ТО } (\textit{Зона} = \textit{Зона}\_\Psi); \tag{18}$$

$$\text{ЕСЛИ } (\textit{Зона} = \Psi) \text{ И } (\textit{Поправка } X = \textit{Поправка } X\_\Psi) \text{ И } \\ (\textit{Поправка } Y = \textit{Поправка } Y\_\Psi) \text{ ТО } (Z = \\ = \textit{Корректировка } Z\_\Psi). \tag{19}$$

Правило (18) характеризует попадание точки с исходными данными $X(x^*, y^*)$ в область определения Зоны Ψ. Выразим коэффициенты влияния $\nu_{x\to\Sigma}$ и $\nu_{y\to\Sigma}$ переменных *Область* X и *Область* Y соответственно на переменную *Зона*:

$$\begin{aligned} \nu_{x\to\Sigma} = (g_{11} + g_{21})(r_x - Base(\Psi)_x), \\ \nu_{y\to\Sigma} = (g_{12} + g_{22})(r_y - Base(\Psi)_y). \end{aligned} \tag{20}$$

Тогда, используя определенный формулой (3) порядок реализации этапа логического вывода, для исходных данных, заданных вектором $\vec{x}_0(x_0, y_0)$, получим

$$\Sigma = \textit{Область } X\_\Psi(x_0) \cdot \nu_{x \to \Sigma} + \textit{Область } Y\_\Psi(y_0) \cdot \nu_{y \to \Sigma} =$$
$$= \frac{x_0}{r_x - Base(\Psi)_x} \cdot (g_{11} + g_{21})(r_x - Base(\Psi)_x) +$$
$$+ \frac{y_0}{r_y - Base(\Psi)_y} \cdot (g_{12} + g_{22})(r_y - Base(\Psi)_y) =$$
$$= x_0(g_{11} + g_{21}) + y_0(g_{12} + g_{22}),$$

что совпадает с формулой (16), выражающей условие принадлежности точки исходных данных области определения Зоны.

Выразим коэффициенты влияния $\nu_{\textit{Зона} \to Z}$,$\nu_{x \to z}$ и $\nu_{y \to z}$ переменных *Зона*, *Поправка* X и *Поправка* Y соответственно на переменную Z в виде

$$\begin{aligned} \nu_{\textit{Зона} \to Z} &= Base(\Psi)_z; \\ \nu_{x \to z} &= (g_{11}a_\perp + g_{21}b_\perp)(r_x - Base(\Psi)_x); \\ \nu_{y \to z} &= (g_{12}a_\perp + g_{22}b_\perp)(r_y - Base(\Psi)_y). \end{aligned} \tag{21}$$

Тогда, используя порядок реализации этапа логического вывода по формуле (3), получим значение выводимой переменной Z:

$$Z = \textit{Зона}\_\Psi(\Sigma) \cdot Base(\Psi)_z + \textit{Поправка } X\_\Psi(x_0) \cdot (g_{11}a_\perp + g_{21}b_\perp) \times$$
$$\times (r_x - Base(\Psi)_x) + \textit{Поправка } Y\_\Psi(y_0) \cdot (g_{12}a_\perp + g_{22}b_\perp) \times$$
$$\times (r_y - Base(\Psi)_y) =$$
$$= 1 \cdot z_0 + \frac{x_0}{(r_x - Base(\Psi)_x)}(g_{11}a_\perp + g_{21}b_\perp)(r_x - Base(\Psi)_x) +$$
$$+ \frac{y_0}{(r_y - Base(\Psi)_y)}(g_{12}a_\perp + g_{22}b_\perp)(r_y - Base(\Psi)_y) =$$
$$= z_0 + (g_{11}a_\perp + g_{21}b_\perp)x_0 + (g_{12}a_\perp + g_{22}b_\perp)y_0,$$

что совпадает с выражением (15), которое определяет величину выходной переменной, соответствующей положению гиперплоскости Зоны Ψ. Полученное значение Z будет использовано на этапе композиции с ненулевым весовым коэффициентом.

Дополняя имеющийся набор корректировочных данных новой точкой X_{New}, необходимо установить, попадает ли она в одну из имеющихся областей $Def(\Psi_i)$. Если такая область отсутствует, то необходимо построить одну новую Зону с участием X_{New} без нарушения условия взаимного пересечения областей $Def(\Psi_i)$. Если X_{New} попадает в одну из областей $Def(\Psi_J)$, то следует образовать $(N+1)$ новую Зону вместо Зоны Ψ_J, используя для формирования базиса зоны точку X_{New} в качестве основания Зоны и точки, входящие в построение зоны Ψ_J.

Предложенная стратегия построения Зон позволит получать решение для выводимой переменной, которое принадлежит гиперплоскости, аппроксимирующей закономерность взаимосвязи переменных на сколь угодно малых областях $Def(\Psi_i)$, что обеспечит приближение к произвольной непрерывной функции с точностью до любого $\varepsilon > 0$.

Таким образом, рассмотренная модификация нечеткого вывода позволяет снизить влияние субъективного фактора, ухудшающего качество решения вследствие неточностей, которые вносит эксперт при описании системы. При решении задач управления в областях исходных данных, характеризуемых неполнотой описания функционирования системы, благодаря рассмотренному подходу с помощью набора корректировочных данных удается построить решение практически с требуемой точностью в тех областях, где знания эксперта оказываются неточными или ошибочными. При этом сохраняется объяснительная возможность нечеткого вывода решения в терминах, введенных экспертом.

СПИСОК ЛИТЕРАТУРЫ

1. W a n g L., M e n d e l J. M. Fuzzy basis functions, universal approximation, and orthogonal least-squares learning // IEEE Transactions Neural Networks. – September 1992. – Vol. 3, No. 5. – P. 807–814.
2. K o s k o B. Fuzzy systems as universal approximators // IEEE Transactions on Computers. – November 1994. – Vol. 43, No. 11. – P. 1329–1333.
3. C a s t r o J. L., D e l g a d o M. Fuzzy systems with *Def* uzzification are universal approximators // IEEE Transactions on System, Man, and Cybernetics. – April 1995. – Vol. 25, No. 4. – P. 629–635.
4. T s u k a m o t o T. An Approach to Fuzzy Reasoning Method // Advances in Fuzzy Set Theory and Applications. – 1979. – P. 137–149.
5. T a k a g i T., S u g e n o M. Fuzzy identification of systems and its applications to modeling and control // IEEE Transactions on System, Man, and Cybernetics. – 1985. – Vol. 15, No. 1.– P. 116–132.
6. К р у г л о в В. В., Д л и М. И. Нечеткая логика и искусственные нейронные сети. – М.: ФИЗМАТЛИТ, 2001. – 224 с.

Статья поступила в редакцию 10.05.2012