

ВИЗУАЛИЗАЦИЯ ПРОЦЕССОВ ПРИНЯТИЯ РЕШЕНИЙ

О. Г. Берестнева, В. А. Воловоденко

Институт кибернетики Национального исследовательского
Томского политехнического университета, 634034, Томск, Россия

УДК 004.9

Рассмотрены проблемы визуализации многомерных нечетких рядов. Предложена методика анализа образов основы нечеткого ряда. Показано, что наиболее жесткой конструкцией нечеткого ряда является компонента ряда с наибольшим значением функции принадлежности. Построенный в соответствии с предлагаемой методикой ряд визуальных образов отражает тенденцию развития нечеткого временного ряда. Применение методов визуализации нечетких временных рядов способствует выявлению качественных связей при рассмотрении динамических слабоструктурированных систем (в том числе социальных и экономических). Предложенный подход позволяет на качественном уровне решать задачи группирования, структурирования и прогнозирования в таких системах.

Ключевые слова: нечеткий временной ряд, слабоструктурированные системы, принятие решений, визуализация многомерных данных.

The problems of visualization of multidimensional fuzzy ranks. The methodology of analysis of images basis of fuzzy number. It is shown that the most rigid structure of a fuzzy number is the number of components with the highest value of membership function. Built in accordance with the proposed method a series of visual images reflect the tendency of the fuzzy time series. Application of fuzzy time series visualization facilitates the identification of high-quality bonds in the consideration of dynamical semi-system (including social and economic). The proposed approach allows for a qualitative level to solve the problem of grouping, structuring and forecasting in such systems.

Key words: Fuzzy time series, semistructured systems, decision-making, visualization of multidimensional data.

Введение. Основной особенностью систем принятия решений является усиливающаяся математизация таких систем. Применение математических методов обусловлено возможностью использования однозначных процедур, которые могут выполняться на компьютерах. При этом большинство решаемых на компьютере задач имеют явный алгоритмический характер, требуют существования решения и корректных вычислительных методов. Классическое решение един-

ственно и достоверно, достижимо за конечное, практически приемлемое время с использованием существующих методов поиска.

Использование методов принятия решений в социальных и экономических системах свидетельствует о том, что во многих случаях данные условия не могут быть выполнены. Существование решения можно доказать только в отдельных случаях. При переходе от практической постановки задачи к ее математическому аналогу не учитывается большое количество влияющих на решение факторов, анализ практической постановки задачи позволяет сделать вывод о возможности существования множества "решений". В таких случаях лицо, принимающее решение, оказывается в ситуации, когда выбор решения превращается в процесс выбора среди множества альтернатив, часто основанных на противоречивых условиях. Следует учитывать, что в социальных и экономических системах действует большое количество факторов, которые изменяются во времени. При отслеживании состояния таких систем возникают проблемы временных рядов. Эти ряды обладают высокой степенью неопределенности, для их обработки необходимо применять методы, не обеспечивающие высокую точность результата. Примером могут служить нейросетевые и нечеткие модели, а также модели искусственного интеллекта.

Моделирование поведения социальных и экономических систем основано на модели нечеткого динамического процесса, получившей название нечеткого временного ряда (НВР). Многомерность НВР – свойство, обеспечивающее комплексность представления информации при сохранении независимости между фиксируемыми показателями.

В данной работе рассматривается простейший случай, когда многомерный нечеткий временной ряд (МНВР) представлен показателями, которые поддаются измерению и выражаются вещественными числами:

$$\mathbf{W} = \{ \mathbf{V}(t_1), \mathbf{V}(t_2), \dots, \mathbf{V}(t_i), \dots, \mathbf{V}(t_k) \}. \quad (1)$$

Здесь \mathbf{W} – МНВР, конечный набор векторов показателей; $t_1, t_2, \dots, t_i, \dots, t_k$ – моменты времени, когда наблюдается и регистрируется $V(t_i)$:

$$\mathbf{V}(t_i) = \{ q_1(t_i), q_2(t_i), \dots, q_n(t_i) \}. \quad (2)$$

Вектор показателей измеряется и регистрируется в момент t_i . Нечеткость скрыта в показателях $q_j(t_i)$.

Анализ МНВР затруднен тем, что получение полных количественных данных не представляется возможным.

Постановка задачи. Для известного МНВР требуется построить визуальный образ, позволяющий ЛПР проводить необходимые манипуляции и анализ МНВР.

Из представлений (1), (2) следует, что многомерный нечеткий временной ряд \mathbf{W} проще представить в виде электронной таблицы или списка списков. В этом случае $\mathbf{V}(t_i)$ – вектор-строка, состоящая из различных показателей, измеряемых и регистрируемых в момент t_i :

$$\mathbf{V}(t_i) = \{ q_1(t_i), q_2(t_i), \dots, q_n(t_i) \} = \{ q_{i_1}, q_{i_2}, \dots, q_{i_n} \}.$$

Здесь $q_{i_n} \in R$; $V(t_i)$ можно считать точкой пространства R_n для момента t_i .

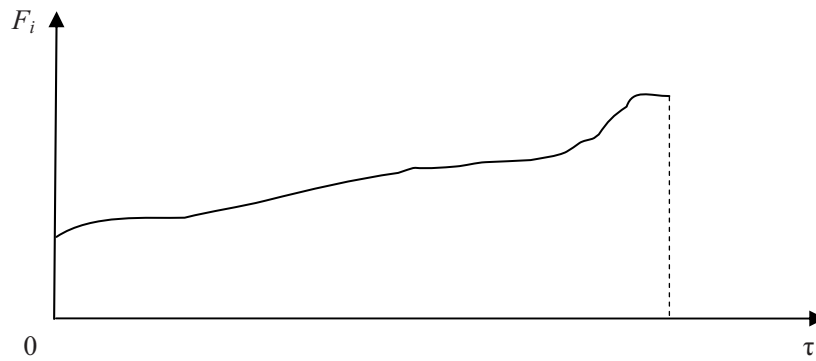


График функции $F_i(\tau)$

Из (1) следует, что МНВР представляет собой последовательность точек R_n :

$$\mathbf{W} = \{\mathbf{V}(t_1), \mathbf{V}(t_2), \dots, \mathbf{V}(t_i), \dots, \mathbf{V}(t_k)\} = \{\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_i, \dots, \mathbf{V}_k\}.$$

Это конечная последовательность, которую можно считать отрезком МНВР.

Введем в рассмотрение множество ортонормированных на $[0, 1]$ полиномов

$$\mathbf{P} = \{P_0(\tau), P_1(\tau), \dots, P_{n-1}(\tau)\}^T = \text{column}\{P_0(\tau), \dots, P_{n-1}(\tau)\}.$$

Основной особенностью этих полиномов будем считать то обстоятельство, что их можно представить в виде графиков функций аргумента $\tau \in [0, 1]$. Тогда каждой строке $\mathbf{V}(t_i)$ можно поставить в соответствие функцию

$$F_i(\tau) = (\mathbf{V}(t_i), \mathbf{P}) = \sum q_{i,k} P_k(\tau). \tag{3}$$

Последовательности \mathbf{W} поставим в соответствие последовательность $\{F_i(\tau)\}_{i=1}^{i=k}$.

В формуле (3) переменная τ называется параметром композиции, так как способствует переходу от вектора $\mathbf{V}(t_i)$ к функции $F_i(\tau)$ и тем самым порождает визуальный образ $F_i(\tau)$ вектора $\mathbf{V}(t_i)$. Последний можно наблюдать только как набор значений в электронной таблице, а функцию $F_i(\tau)$ можно представить в виде графика (см. рисунок).

Визуализация нечетких временных рядов. Поскольку визуальные свойства ортонормированных полиномов в наибольшей степени проявляются на отрезке $[0, 1]$, он является основой для представления образов $F_i(\tau)$ и прообразов $\mathbf{V}(t_i)$ [1, 2]. Возникает два пространства: пространство оригиналов $\mathbf{V}(t_i)$ и пространство изображений-образов. Два эти пространства связаны между собой условием изометричности. Действительно, евклидова норма в пространстве оригиналов R_n определяется по формуле

$$\|\mathbf{V}(t_i)\|_{R_n} = \sqrt{\sum_{k=1}^n q_{ik}^2},$$

а норма функции $F_i(\tau)$ в пространстве изображений $L_n(\tau)$ может быть вычислена по формуле

$$\|F_i(\tau)\|_{L_n} = \left\| \sum_{k=1}^n q_{ik} P_{k-1}(\tau) \right\|_{L_n} = \left(\int_0^1 \left(\sum_{k=1}^n (q_{ik} P_{k-1}(\tau))^2 \right) d\tau \right)^{\frac{1}{2}} = \left(\sum_{k=1}^n q_{ik}^2 \right)^{\frac{1}{2}}. \tag{4}$$

Последнее равенство выполняется в силу свойств ортонормирования полиномов $P(\tau)$. Таким образом, выполняется равенство $\|V(t_i)\|_{R_n} = \|F_i(\tau)\|_{L_n}$, из которого следует изометричность двух пространств.

Через функцию нормы можно определить функцию метрики r

$$r(A(\tau), B(\tau)) = \|A(\tau) - B(\tau)\|$$

и различия между двумя функциями – образами $A(\tau)$ и $B(\tau)$.

Соответствие $A(t_i) \rightarrow A(\tau)$, $B(t_j) \rightarrow B(\tau)$ позволяет наблюдать различие образов на экране компьютера. При этом происходит включение аналитических способностей ЛППР в схему исследования образов, а значит, и оригиналов. Зрительная система человека является наиболее быстрой системой анализа изображений и позволяет делать выводы о свойствах наборов оригиналов. Таким образом, происходит разделение функций между ЛППР и компьютером. На компьютере осуществляются рутинная обработка электронной таблицы и подготовка образов, а ЛППР выполняет более "тонкую", аналитическую обработку данных.

Если трактовать $A(t_i)$ и $B(t_j)$ как две точки многомерного пространства, то отрезок AB можно представить в виде

$$AB = (1 - \lambda)A(t_i) + \lambda B(t_j), \quad \lambda \in [0, 1].$$

Параметр λ называется секвенциальным параметром, т. е. параметром, определяющим последовательность образов.

На основе трех переменных (параметра композиции τ , переменной для значения образов F , секвенциального параметра λ) можно сформировать визуальное пространство $\{\tau, F(\tau, \lambda), \lambda\}$, которое используется для представления МНВР.

В качестве параметра λ можно использовать функцию метрики r . В этом случае получается наглядная картина представления МНВР на экране компьютера.

Секвенциальный параметр λ может играть роль времени t . Заметим, что выполнен переход от последовательности оригиналов к последовательности визуальных образов, у которых сохранены все информационные признаки нечеткости. Признаки нечеткости могут быть сохранены в электронной таблице оригиналов как значения функции принадлежности каждого показателя.

С практической точки зрения все оригиналы принадлежат некоторому универсуму, при этом каждый показатель ассоциируется с некоторым свойством, способным порождать нечеткий предикат, который можно характеризовать значением истинности. Принято считать, что нечеткий предикат может принимать континуум значений истинности $[0, 1]$, причем число 0 соответствует понятию "нуль", а число 1 – понятию "истина". Чем в большей степени показатель соответствует рассматриваемому свойству, тем более близким к единице должно быть значение истинности нечеткого предиката [3].

Функции принадлежности можно использовать в качестве параметров, которые определяют секвенциальный параметр λ . Следовательно, визуальные образы $F_i(\tau)$ будут упорядочены в соответствии со значениями функций принадлежности. Для всех образов $F_i(\tau)$ фиксируется одна функция принадлежности с разными значениями для каждого образа в соответствии с данными.

Прием визуализации позволяет исследовать влияние конкретной функции принадлежности на порядок следования образов и их расположение. Эту информацию можно получить в виде списков, однако визуальная картина охватывает весь ансамбль образов и позволяет выделить в нем наиболее "выдающиеся" образы. Эти образы будут группироваться ближе к концам отрезка секвенциальности, а весь ансамбль образов можно характеризовать точками сгущения образов. Вид функции принадлежности будет влиять на последовательность образов, поэтому функцию принадлежности можно задавать экспертным путем, ориентируясь на такие свойства, которые могут быть измерены по некоторой известной количественной шкале. Так как НВР не требует точного задания функций принадлежности, иногда достаточно зафиксировать наиболее характерные значения [4].

Визуальное группирование данных позволяет ЛПР представить структуру НВР и определить тенденцию развития этого ряда. Однако при формировании визуальных образов надо учитывать, что показатель $q_k(t)$ может быть представлен в виде списка нечетких величин μ_{ki} . Это означает, что

$$q_k(t) = \{ \{q_{k1}(t), \mu_{k1}(t)\}, \{q_{k1}(t), \mu_{k1}(t)\}, \dots, \{q_{ki}(t), \mu_{ki}(t)\} \}.$$

Здесь $q_{ki}(t)$ – значение показателя, для которого определено значение функции принадлежности, ставящей в соответствие некоторое действительное число из интервала $[0,1]$. Таким образом, показатель $q_k(t)$ представлен нечетким множеством.

Очевидно, что при таком обилии значений визуализация, т. е. переход от оригиналов к образам, представляет собой проблему. Для решения этой задачи можно предложить несколько подходов, один из которых состоит в следующем: из данных, представленных НВР \mathbf{W} , выделяется наиболее достоверный НВР. Этого можно достичь, если от \mathbf{W} перейти к НВР W_0 , который в качестве $q_k(t)$ имеет не множество, а одно нечеткое значение, но это значение $q_k^0(t)$ является наиболее достоверным, а значит, имеет максимальное значение функции принадлежности в множестве $q_k(t)$. Таким образом, получаем

$$q_k^0(t) = \max_{\mu_{ki}} \{q_{ki}, \mu_{ki}\}.$$

Следовательно,

$$W_0 = \{V^0(t_1), V^0(t_2), \dots, V^0(t_k)\}, \quad V^0 = \{q_1^0(t_i), q_2^0(t_i), \dots, q_{ki}^0(t_i)\}.$$

Нечеткий временной ряд W_0 называется нулевым базовым временным рядом, для которого можно построить множество $\{F_i^0(\tau)\}_{i=1}^{i=k}$ на основе изложенных выше принципов. В результате получаем визуальный образ для W_0 , который обозначим F_0 . Анализ образа F_0 показывает, что это нечеткий образ и его свойства являются следствием свойств W_0 . В то же время возможна обратная связь, т. е. по свойствам F_0 можно судить о свойствах W_0 .

При достаточно последовательном развитии временного ряда его изменения не являются существенными и получаемый образ F_0 обладает свойствами W_0 . Можно считать W_0 жестким "каркасом" \mathbf{W} , соответственно F_0 будет более устойчивой конструкцией, чем F . При визуализации устойчивость конструкции будет проявляться в том, что при изменении \mathbf{W} нулевой базо-

вый временной ряд W_0 не будет меняться, если при изменении \mathbf{W} компонент W_0 останется неизменным. Аналогичные рассуждения можно провести для образа F_0 .

Если компоненты W_0 исключить из \mathbf{W} , то можно получить W_1 – первый базовый временной ряд и поставить ему в соответствие $\{F_i^1(i)\} = F_1$. Образ F_1 менее устойчив к изменениям \mathbf{W} и показывает более вариабельную картину. При анализе изображения F_1 можно установить более широкие границы области изменения образов $\{F_i^1(\tau)\}$. В силу линейности преобразования (4) сохраняется аддитивный характер переходов F_0, F_1, \dots, F_{n-1} .

Заключение. На основе визуализации возможно решение ряда задач, таких как сегментация, кластеризация, прогнозирование, для которых затем необходимо обнаружить аналогичную форму постановки. В решении подобных задач заинтересованы прежде всего различные ЛПР, для которых будет упрощен процесс формирования результатов.

Заметим, что изложенный подход не связан с существующими математическими моделями системы, для которой ведется визуальный анализ НВР. Это обуславливает высокую точность прогнозирования.

Список литературы

1. БЕРЕСТНЕВА О. Г., ВОЛОВОДЕНКО В. А., ШАРОПИН К. А. Визуализация экспериментальных многомерных данных на основе обобщенных графических образов // Вестн. науки Сибири. Сер. Информ. технологии и системы управления. 2011. №. 1. С. 363–369. [Электрон. ресурс]. <http://sjs.tpu.ru/journal/article/view/75>.
2. VOLOVODENKO V. A., BERESTNEVA O. G., SHAROPIN K. A. Visual interpretation of quantitative characteristics of biosystems // Fundamental medicine: from scalpel toward genome, proteome and lipidome: Proc. of the 1st Intern. conf., Kazan, Apr. 25–29, 2011. Kazan: KSU, 2011. P. 126–129.
3. Нечеткие множества / Под ред. Д. А. Поспелова. М.: Наука, 1986.
4. ЛЕОНЕНКОВ А. В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. СПб.: БХВ-Петербург, 2005.

Берестнева Ольга Григорьевна – д-р техн. наук, проф. Института кибернетики Томского политехнического университета; тел.: (382-2) 42-61-00; e-mail: ogb@tpu.ru;
Воловоденко Виталий Алексеевич – канд. техн. наук, доц. Института кибернетики Томского политехнического университета; тел.: 8 (382-2) 42-13-11; e-mail: volcowvav@tpu.ru

Дата поступления – 21.09.12 г.