Д. А. Молодцов, А. В. Осин, Новый метод применения многозначных закономерностей,
Нечеткие системы и мягкие вычисления, 2020, том 15, выпуск 2, 83–95

УДК 519.7

НОВЫЙ МЕТОД ПРИМЕНЕНИЯ МНОГОЗНАЧНЫХ ЗАКОНОМЕРНОСТЕЙ[1]

Молодцов Д.А.*, Осин А.В.**

*Вычислительный центр им. А.А. Дородницына Российской академии наук Федерального исследовательского центра «Информатика и управление» Российской академии наук, г. Москва

**Московский технический университет связи и информатики, г. Москва

Поступила в редакцию 24.06.2020, после переработки 13.07.2020.

В работе предлагается метод построения многозначных закономерностей, основанный на сравнении закономерностей для текущего значения аргументов. Приведено подробное описание всех алгоритмов. Эффективность метода продемонстрирована на примерах прогнозирования некоторых финансовых индексов.

Ключевые слова: многозначная закономерность, портфельный подход.

Нечеткие системы и мягкие вычисления. 2020. Т. 15, № 2. С. 83–95.
https://doi.org/10.26456/fssc72

Введение

Поиск и нахождение закономерностей является важнейшей частью любой науки. Наиболее распространенный способ описание закономерностей на языке математики – это использование однозначной функции $f : X \to Y$, значениями которой являются значения описываемой переменной. При практической проверке такой закономерности могут возникнуть ситуации, когда при одном и том же значении аргумента появляются результаты опытов с различными значениями функции. Для решения этой проблемы зачастую предполагают, что в закономерности присутствует случайная компонента с определёнными характеристиками, что и объясняет разброс реальных значений. Иногда такой подход приемлем, но так происходит далеко не всегда. Если изначально список аргументов закономерности был недостаточным, то списывать разброс в опытах на наличие случайности уже неправомерно. Однако заранее узнать достаточный список переменных для построения однозначной закономерности весьма непросто. Таким образом можно сделать вывод, что применение однозначных закономерностей основано на принятии ряда гипотез. Проверка таких гипотез также является непростой задачей и часто вообще не производится.

В работе [1], был предложен другой метод построения закономерностей, свободный от необходимости принятия гипотезы о случайной составляющей. Суть метода заключалась в том, что закономерность описывалась не однозначной функцией, а многозначным отображением в форме мультимножества. Поэтому

[1]Работа выполнена при частичной финансовой поддержке РФФИ (грант №19-01-00625).

проблема построения закономерности практически отпадала. Любой набор результатов опытов в виде таблицы $D_n = \{(x_1 y_1), ..., (x_n y_n)\,|(x_i, y_i) \in X \times Y$ можно рассматривать, как график точечно множественного отображения. Теперь выбор набора аргументов и множества результатов опытов, по сути, означал построение многозначной закономерности. Поэтому основной вопрос при таком подходе состоит не в построении закономерности, а в выборе из имеющихся закономерностей наиболее приемлемой с нашей точки зрения. В [1] была предложена методология сравнения многозначных закономерностей, которая использовалась для выбора одной закономерности и дальнейшего её использования с дополнительными методами экстраполяции многозначных отображений [2,3]. Такая методология показала достаточную эффективность в многих практических применениях. В настоящей работе предлагается ещё один способ применения многозначных закономерностей. Чтобы описать его основные идеи напомним основные моменты метода из работы [1].

1. Методология сравнения многозначных закономерностей

Будем предполагать, что исходным материалом для построения закономерности служит набор результатов опытов $D_n = \{(x_1 y_1), ..., (x_n y_n)\,|(x_i, y_i) \in X \times Y$, где $X \subseteq X_1 \times \ldots \times X_m$. Закономерность, описывающая зависимость переменной y от набора аргументов, определяется указанием непустого подмножества аргументов $I \subseteq \{1, \ldots, m$ и представляет собой мультимножество $X \times Y I \mu$, где $\mu : X(I) \times Y \to N \cup \{0\}$, $X(I) = \prod\limits_{i \in I} X_i$, N – множество натуральных чисел. Значение $\mu(xy)$ показывает сколько раз пара (x, y), $x = \prod\limits_{i \in I} x_i \in X(I)$, $y \in Y$, встречалась в данных D_n. Это значение трактуется, как надежность точки графика (x, y) закономерности $X \times Y I \mu$.

Надежность закономерности при значении аргумента $x \in X(I)$ считается равной сумме $\sum\limits_{y \in Y} \mu(x, y)$. Надежность всей закономерности принимается равной величине

$$Saf\,[X \times Y, I, \mu] = \sum_{y \in Y} \mu\,(x, y)\,.$$

Точность закономерности $X \times Y I \mu$ в точке $x \in X(I)$ называется мощность множества $\{y \in Y\,|\,\mu\,(x, y) > 0\}$, которую будем обозначать символом $|\{y \in Y\,|\,\mu\,(x, y) > 0\}|$. Точность всей закономерности $X \times Y I \mu$ считается равной

$$Ac\,(X \times Y, I, \mu\,) = |\{y \in Y\,|\,\mu\,(x, y) > 0\}|\,.$$

Важное замечание. Критерии надежности и точности никак не используют специфику множеств X и Y и поэтому значения функции и значения аргументов могут иметь произвольную природу.

При построении методологии сравнения представляется естественным сначала отобрать закономерности, достигающие заданного уровня надежности, а потом оставшиеся закономерности сравнивать по точности. Однако, такой способ сравнения имеет один недостаток. Для фиксированной таблицы данных закономерности с большим количеством аргументов будут иметь, вообще говоря, меньшую надежность, чем закономерности с небольшим количеством аргументов. В [1] эта проблема подробно описана. Чтобы решить эту проблему в [1] предложено, если закономерность имеет недостаточный уровень надежности, то применять к ней процедуру замены множества значений аргументов на сеточное множество с меньшим количеством значений. Такая процедура даст увеличение надежности закономерности, но может понизить ее точность. Для того чтобы такая процедура проводилась плавно, с постепенным изменением надежности, нужно задать семейство таких сеток на множестве значений аргументов.

После достижения заданного уровня надежности закономерность сравнивается с другими по точности, но не исходной точности, а точности, соответствующей достигнутому уровню надежности. Если достигнуть нужного уровня надежности не удается, то закономерность признается неудовлетворительной и в сравнении больше не участвует.

После нахождения наилучшей или нескольких наилучших по точности закономерностей их можно использовать для различных целей. Отметим, что в приведенном методе выбора закономерностей не делается никаких предварительных гипотез о том какому классу принадлежит закономерность. Это обстоятельство дает возможность находить любые нелинейные и неоднозначные закономерности без внесения в процесс выбора неконтролируемого произвола.

Найденные закономерности представляются в виде графиков, заданных конечным набором точек. Для дальнейшего использования, например в прогнозировании, такое представление нужно дополнять экстраполяцией. Различные варианты экстраполяции для статического и динамического случая предложены в [2,3]. Вообще говоря, методы экстраполяции закономерностей вносят некоторый произвол в исходную закономерность.

Итак, применение описанного метода для прогнозирования предполагает следующие этапы:

1. Выбор закономерностей.

2. Получение значений аргументов для построения прогноза.

3. Построение экстраполяций для закономерностей для заданных значений аргументов.

4. Построение результирующего прогноза на базе прогнозов по отдельным закономерностям.

Перейдем теперь к описанию нового метода.

2. Альтернативный метод применения многозначных закономерностей

Если проанализировать описанный выше метод применения многозначных закономерностей, то можно отметить, что выбираются закономерности так, чтобы они давали нужную надежность и точность при всех значениях аргумента, а используются эти закономерности всякий раз при одном конкретном значении аргумента. Это наводит на простую мысль, а нельзя ли поменять этапы 1 и 2 в предложенной выше схеме. Сначала получить значение вектора аргумента, а потом уже выбрать закономерности, которые будут наилучшими не для всех значений аргументов, а только для того значения, которое сейчас требуется обработать.

Реализация этой идеи приводит к следующей последовательности этапов построения прогноза.

1. Формирование таблицы исходных данных.

2. Получение новых значений исходной информации и перевод их в таблицу исходных данных.

3. Выбор наилучших многозначных закономерностей.

4. Построение результирующего значения прогноза.

Отметим, что этапа экстраполяции в этой схеме нет. Хотя названия этапов остались в основном прежними, но их содержание существенно изменилось, поэтому перейдем к подробному описанию новых этапов прогнозирования.

2.1. Формирование таблицы исходных данных

Таблица исходных данных состоит из столбца функции и столбцов аргументов. Каждая строка аргументов содержит информацию о значениях аргументов, которые известны к моменту времени, соответствующему этой строке. Строки упорядочены по возрастанию моментов времени. Значения функции также соответствуют моментам времени строк, но с некоторым сдвигом, описываемым целочисленным параметром Slip. Если рассматриваются дневные данные, то сдвиг Slip=1 означает, что сегодняшним значениям аргументов соответствуют завтрашние значения функции. Построение закономерности по такой таблице будет давать прогноз завтрашним значениям функции, используя сегодняшние значения аргументов и все значения функции и аргументов, которые известны к текущему моменту времени. Предполагается, что к моменту времени, соответствующему строке, нам известны и значения аргументов и значения функции. Учитывая описанную структуру таблицы данных, это предположение означает, что к моменту времени, соответствующему строке с номером N, нам известны все значения аргументов на строках с номерами меньших или равных N, и все значения функции на строках с номерами меньшими или равными N-Slip. Такая структура таблицы позволяет строить прогнозы и на 1 день и на 30 дней, если это позволяют объемы данных.

Формирование столбцов аргументов целиком зависит от наших желаний и целей исследований. Это могут быть значения любых параметров, которые мы можем измерять - значения скорости их изменения, значения, соответствующие прошлым моментам времени и т. д. Главное, чтобы соблюдалась временная структура исходной таблицы данных.

Отметим, что пока для простоты в работе описывается только работа с числовыми данными. Работа со строковыми значениями полей также возможна, но она требует задания семейств огрубления шкал для каждого поля и пока не рассматривается.

2.2. Получение новых значений исходной информации и перевод их в таблицу исходных данных

Прежде всего отметим, что применение нового метода состоит из двух различных процессов. Первый процесс – это построение прогнозной функции. При этом процессе мы имеем дело с фиксированной таблицей данных, которая целиком заполнена данными.

Второй процесс – это работа прогнозной функции в режиме реального времени. Здесь происходит получение новых данных о значениях аргументов. Формируется новая последняя строка таблицы, куда и помещаются данные о значениях аргументов. Отметим, что таблица оказывается заполненной только частично. Часть последних значений столбца функции остается незаполненными.

2.3. Выбор наилучших многозначных закономерностей

Как уже отмечалось, выбор закономерности – это выбор непустого подмножества в множестве полей аргументов и выбор некоторого способа огрубления шкал полей аргументов. Шкала – это некоторая функция, которая по числовому значению поля выдает некоторое целое число, которое называется координатой этого значения относительно заданной шкалы. В программе, реализующей описываемый метод, использовались два типа шкал. Первый тип шкалы – арифметический, описывается двумя положительными числами: точностью Acc и шагом Step. Если значение D удовлетворяет неравенству $|D| < Acc$, то координата равна нулю, если выполнены неравенства $Acc + (n-1)\,Step \le |D| < Acc + n\,Step$, то координата равна $n\,sign(D)$.

Второй тип шкалы – логарифмический, также описывается двумя положительными числами: точностью Acc и шагом Step. Если значение D удовлетворяет неравенству $|D| < Acc$, то координата равна нулю, если выполнены неравенства $Acc(1 + Step)^{n-1} \le |D| < Acc\,(1 + Step)^n$, то координата равна $n\,sign(D)$.

Для работы метода нужно задать рассматриваемое множество закономерностей. Если количество полей аргументов равно M, то количество непустых подмножеств равно $2^M - 1$. При небольших значения M можно перебирать все подмножества, при больших значениях перебор нужно ограничить. Поэтому для работы алгоритма нужно указать возможные значения для подмножеств множества полей аргументов. Для краткости эти подмножества будут называться масками закономерности, а множество возможных масок обозначим $Mask$.

Входной информацией для алгоритма является номер строки таблицы - Run, для которой необходимо сделать прогноз. Для каждой маски $M \in Mask$ вычисляется вектор значений аргументов $p\,(Run, M) \in X\,(M)$. Вектор $p(RunM)$ будем называть паттерном. Далее задается горизонт для вычисления статистики паттернов. Обозначим его Gor. Горизонт является параметром метода прогнозирования.

Статистика считается следующим образом. Рассматриваются строки с номерами из интервала $[Run - Gor, Run - Slip]$ Отметим, что для строк с номерами большими чем $Run - Slip$ значения функции, с точки зрения строки Run, еще не известны. Для каждой допустимой маски формируется статистический список $Stat(RunM)$. Для каждого допустимого значения строки t проверяется условие $p\,(t, M) \cong p(RunM)$. Под знаком $\cong$ понимается совпадение соответствующих координат относительно шкал полей аргументов. Если это условие выполнено, то в список добавляется тройка $(tf\,(t)\,W(t))$, где $f\,(t)$ – значение поля функции, $W\,(t)$ – вес этого значения, который вычисляется по формуле

$$W\,(t) = \delta^{Run-t}.$$

Здесь $\delta,\ 0 < \delta \le 1$, коэффициент дисконтирования, который позволяет учесть удаленность значения статистики от строки прогнозирования. Этот коэффициент является параметром метода.

Далее вычисляется надежность паттерна $p(RunM)$, которая равна следующей величине

$$S\left(Run, M\right) = \sum_{(t,f(t))\in Stat(Run,M)} \gamma^{Run-t}$$

Здесь $\gamma,\ 0 < \gamma \leq 1$ – это коэффициент дисконтирования, который дает возможность учитывать совпадения паттернов в зависимости от удаленности от момента времени прогнозирования. Коэффициент дисконтирования γ – это тоже параметр метода.

Задается еще один параметр метода – уровень надежности S. Если $S\left(Run, M\right) \geq S$, то данный шаблон участвует в прогнозировании. Если неравенство не выполнено, то выполняется последовательное синхронное огрубление всех полей аргументов, входящих в маску, до тех пор, пока ограничение по надёжности не будет выполнено. Если в результате процесса огрубления ограничение по надёжности так и не выполнено, то данный шаблон не участвует в построении прогноза.

Огрубление производится с помощью некоторого оператора, который называется трансформацией, на вход трансформации подается шкала $(AccStep)$, а на выходе получается огрубленная шкала $(NewAccNewStep)$. Этот оператор описывается двумя операциями $\alpha\beta$ и тремя числами DeltaAcc, DeltaStep, K. Операции $\alpha\beta$ могут быть сложением, умножением или отсутствием операции. Для примера приведем работу трансформации для сложения

$$NewAcc = Acc + K\ DeltaAcc \quad NewStep = Step + K\ DeltaStep.$$

Отметим, что описанный способ отбора масок существенно отличается от способа отбора, предложенного в [1]. Здесь используется надежность только одного значения закономерности, именно того значения, которое сейчас реализовалось и может использоваться для прогнозирования. В [1] учитывались надежности всех значений закономерности.

Обозначим множество масок, оставшееся после такого отбора, символом $SMask(Run)$.

2.4. Построение результирующего значения прогноза

После окончания отбора паттернов по надежности производится отбор паттернов по точности. Считается максимальный разброс значений прогноза в каждом паттерне и оставляются паттерны с минимальным и близким к минимальному разбросам. Процент близости задается специальным параметром метода Range Percent. После этого отбора у нас имеется множество паттернов со статистикой

$$\left\{p\left(Run, M\right), Stat\left(Run, M\right),\ M \in SMask(Run)\right\}$$

Для каждого паттерна теперь считается статистика появления значений поля функции $u \in F$, но складываются не факты появления этих значений, а их веса с учетом дисконтирующего коэффициента δ, $0 < \delta \leq 1$, по формуле

$$Q\left(Run, M, u\right) = \sum_{\substack{(t, f\left(t\right)) \in Stat\left(Run, M\right) \\ f\left(t\right) = u}} \delta^{Run-t}$$

По сути, функция $Q\,(Run, M, u)$ – это аналог функции кратности при описании многозначной закономерности, но с учетом удаленности по времени от текущего момента времени.

Затем производится нормировка по формулам

$$q\,(Run, M, u) = \frac{Q\,(Run, M, u)}{\sum\limits_{u \in F} Q\,(Run, M, u)}$$

Далее статистики $q\,(Run, M, u)$ обрабатываются с целью выделения наиболее "сильных" значений. Для этого оставляются только те $q\,(Run, M, u)$ которые реализуют максимум с точностью до параметра $Soft\,Percent$, $0 \leq Soft\,Percent \leq 1$. Вычисление производится по формуле

$$Strong\,(Run, M) = \{u \in F \mid q\,(Run, M, u) \geq (Soft\,Percent)/100q\,(Run, M, v)\}$$

Остальные статистики отбрасываются.

Теперь все готово для выдачи окончательного значения прогноза. Вид окончательного значения прогноза можно выбирать различными способами. Можно просто выдавать всю полученную статистику, можно формировать мягкое число, как прогнозное значение, можно выдавать множество значений с весами, превышающими заданный уровень и т.д. В программе выбрана простейшая форма прогноза. Прогноз выдается, как число, равное математическому ожиданию значений в статистике паттернов, где роль вероятности этих значений играют подсчитанные суммарные веса значений.

2.5. Симуляция работы алгоритма на интервале

Для оценки работы алгоритма на заданном интервале строк было реализовано последовательное выполнение алгоритма с вычислением критериев эффективности работы алгоритма.

В качестве таких критериев были выбраны следующие:

- Количество шагов с отказом от прогноза.
- Средняя относительная точность прогноза.
- Максимум относительной точности на интервале.

3. Примеры практического применения алгоритма прогнозирования

3.1. Прогнозирование индекса Dow Jones

Были закачаны дневные значения Индекса Dow Jones с 02.01.1987 по 12.06.2020. В качестве столбца функции были взяты цены CLOSE, в качестве столбцов аргументов были взяты цены OPEN, HIGH, LOW, CLOSE со сдвигами на 1,2 и 3 дня. В результате получилась таблица с 13 столбцами и 9088 строками. Таким образом прогноз строился на 1 день вперед.

Сначала был подсчитан результат прогнозирования с параметрами, которые задаются по умолчанию. Список основных из них следующий:

$S = 2;\ \gamma = 0,99; \quad \delta = 0,95; \quad Gor = 100; \quad Range\,Percent = 10; \quad Soft\,Percent = 10.$

В качестве множества масок брались все комбинации столбцов аргументов, содержащие не более трех столбцов. Таких комбинаций было 298. Все типы шкал для полей аргументов были арифметические.

Симуляция проводилась на интервале 100-400 и на интервале 400-9087. Результаты:

Интервал 100-400

- Средняя относительная точность в процентах –2,0562
- Максимальная относительная точность в процентах –39,7308
- Количество шагов с отказом от прогноза – 0.

Интервал 400-9087

- Средняя относительная точность в процентах –1,2222
- Максимальная относительная точность в процентах –100,0118
- Количество шагов с отказом от прогноза – 9.

Далее на интервале 100-400 была проведена оптимизация по указанным параметрам, кроме параметра Gor. Минимизировалась средняя относительная точность.

Результаты:

Значения параметров

$S = 0,8;\ \gamma = 0,89; \quad \delta = 0,7; \quad Gor = 100; \quad Range\,Percent = 30; \quad Soft\,Percent = 15.$

Значения критериев на интервале 100-400

- Средняя относительная точность в процентах –1,1993
- Максимальная относительная точность в процентах –29,2172
- Количество шагов с отказом от прогноза – 0.

Значения критериев на интервале 400-9087

- Средняя относительная точность в процентах –0,7127
- Максимальная относительная точность в процентах –14,8456
- Количество шагов с отказом от прогноза – 0.

Интересно отметить, что в интервал точности от 0 до 4 процентов попадает 99,5626 процентов предсказаний. Как видим результат проверки на интервале почти в 29 раз длиннее интервала настройки параметров, дает значения критериев почти в 2 раза лучше, чем при оптимизации.

Следующие две симуляции предсказаний проводились на интервале 400-9087 для множества масок, содержащих не более двух комбинаций полей и не более одного поля в комбинации. Все остальные параметры не изменялись.

Результаты для масок из двух полей:

- Средняя относительная точность в процентах –0,7272
- Максимальная относительная точность в процентах –14,8456
- Количество шагов с отказом от прогноза – 0.

Результаты для масок из одного поля:

- Средняя относительная точность в процентах –0,8040
- Максимальная относительная точность в процентах –14,8456
- Количество шагов с отказом от прогноза – 0.

Как видим добавление комбинаций полей в множество масок вносит улучшение в среднюю относительную точность, хотя и не очень большие.

Далее из масок по одному полю были исключены поля со сдвигом по времени на три строки. Результаты симуляции:

- Средняя относительная точность в процентах –0,8253
- Максимальная относительная точность в процентах –14,8456
- Количество шагов с отказом от прогноза – 0.

Затем были оставлены только поля со сдвигом в одну строку. Результаты симуляции

- Средняя относительная точность в процентах –0,9317
- Максимальная относительная точность в процентах –100,0058
- Количество шагов с отказом от прогноза – 8.

Как видим результаты существенно ухудшились, что говорит о полезности добавления в аргументы полей с большими, чем 1, сдвигами по времени. Для дополнительного подтверждения этого тезиса была проведена еще одна симуляция с теми же параметрами и на том же интервале, но с масками мощностью не более 4. Результаты:

- Средняя относительная точность в процентах –0,7098
- Максимальная относительная точность в процентах –14,8456
- Количество шагов с отказом от прогноза – 0.

3.2. Прогнозирование индекса NASDAQ100

Была скачана информация о котировках начиная с 12.08.1991 по 12.06.2020. По аналогии с предыдущим параграфом была создана таблица данных.

Для симуляции были выбраны те же параметры, которые получились после оптимизации в предыдущем параграфе. Множество масок содержало все маски с мощностью, не превышающей 3 поля. Интервал симуляции 100-7236. Результаты:

- Средняя относительная точность в процентах –1,2372
- Максимальная относительная точность в процентах –15,8046
- Количество шагов с отказом от прогноза – 0.

Множество масок содержало все маски с мощностью, не превышающей 2 поля. Интервал симуляции 100-7236. Результаты:

- Средняя относительная точность в процентах –1,2585
- Максимальная относительная точность в процентах –15,8046
- Количество шагов с отказом от прогноза – 0.

Множество масок содержало все маски с мощностью, не превышающей 1 поля. Интервал симуляции 100-7236. Результаты:

- Средняя относительная точность в процентах –1,4191
- Максимальная относительная точность в процентах –15,8046
- Количество шагов с отказом от прогноза – 0.

3.3. Прогнозирование индекса SANP-500

Была скачана информация о котировках начиная с 16.02.2001 по 12.06.2020. По аналогии с предыдущим параграфом была создана таблица данных.

Все параметры остались неизменными, а маски брались мощностью не более 3 полей. Результаты симулирования на интервале 100-5216:

- Средняя относительная точность в процентах –0,7808
- Максимальная относительная точность в процентах –13,6158
- Количество шагов с отказом от прогноза – 0.

Все параметры остались неизменными, а маски брались мощностью не более 2 полей. Результаты симулирования на интервале 100-5216:

- Средняя относительная точность в процентах –0,7878
- Максимальная относительная точность в процентах –13,6158
- Количество шагов с отказом от прогноза – 0.

Все параметры остались неизменными, а маски брались мощностью не более 1 поля. Результаты симулирования на интервале 100-5216:

- Средняя относительная точность в процентах –0,8768
- Максимальная относительная точность в процентах –13,9651
- Количество шагов с отказом от прогноза – 0.

Заключение

Предложена методология построения многозначной закономерности на базе конечного семейства многозначных закономерностей, порождаемых таблицей данных. По сути, этот метод является одной из реализаций портфельного подхода к построению закономерностей. Специфика метода состоит в том, что наиболее подходящие закономерности выбираются не заранее, а для каждого набора значений аргументов.

Использование многозначных закономерностей позволяет отказаться от априорных гипотез о модели данных, что делает метод весьма универсальным.

Параметры метода можно разбить на две основные группы. Первая группа параметров: $S.\gamma \quad \delta \quad Gor$ регулирует отношение к прошлому, какое прошлое брать и с какой скоростью происходит “забывание” прошлых данных. Вторая группа: $Range\,Percent \quad Soft\,Percent$ регулирует выбор наилучших закономерностей. Практические вычисления показывают, что эти параметры являются весьма устойчивыми и их “оптимальность” сохраняется весьма долго.

Несмотря на универсальность метода, его эффективность зависит прежде всего от тех полей, которые были выбраны в качестве аргументов. При неудачном выборе полей никакая оптимизация параметров не даст хороших результатов. Это свойство можно использовать и для оценки влияния отдельных полей на функцию.

Предлагаемый универсальный метод является простым и удобным инструментом для построения многозначных закономерностей. Пока его эффективность проверена только на числовых данных, но идеология метода позволяет строить закономерности и для строковых данных.

Список литературы

[1] Молодцов Д.А. Сравнение и продолжение многозначных зависимостей // Нечеткие системы и мягкие вычисления. 2016. Т. 11, № 2. С. 115–145.

[2] Молодцов Д.А. Экстраполяция многозначных зависимостей // Нечеткие системы и мягкие вычисления. 2017. Т. 12, № 1. С. 45–63.

[3] Молодцов Д.А. Мягкая динамическая экстраполяция многозначных зависимостей // Нечеткие системы и мягкие вычисления. 2019. Т. 14, № 1. С. 5–18.

Образец цитирования

Молодцов Д.А., Осин А.В. Новый метод применения многозначных закономерностей // Нечеткие системы и мягкие вычисления. 2020. Т. 15, № 2. С. 83–95.
https://doi.org/10.26456/fssc72

Сведения об авторах

1. **Молодцов Дмитрий Анатольевич**
ведущий научный сотрудник Вычислительного центра им. А.А. Дородницына Российской академии наук Федерального исследовательского центра «Информатика и управление» Российской академии наук.

Россия, 119311, г. Москва, ул. Вавилова, д. 40, ВЦ РАН.

2. **Осин Андрей Владимирович**
доцент кафедры «Информационная безопасность» Московского технического университета связи и информатики.

Россия, 111024, г. Москва, Авиамоторная улица, дом 8а.

A NEW METHOD OF APPLYING MULTI-VALUED DEPENDENCIES

Molodtsov Dmitriy Anatolievich
Leading Researcher at Institution of Russian Academy of Sciences,
Dorodnicyn Computing Centre of RAS
Russia, 119333, Moscow, 40 Vavilova str., CC RAS.

Osin Andrey Vladimirovich
Associate Professor at Information Security department, Moscow Technical University of Communications and Informatics
Russia, 111024, Moscow, 8a Aviamotornaya str.

Received 24.06.2020, revised 13.07.2020.

The paper proposes a method for constructing multi-valued dependencies based on a comparison of dependencies for the current value of the arguments. A detailed description of all the algorithms is given. The effectiveness of the method is demonstrated by forecasting some financial indices.

Keywords: multi-valued dependencies, portfolio approach in forecasting.

Citation

Molodtsov D.A., Osin A.V., "A new method of applying multi-valued dependencies", *Nechetkie Sistemy i Myagkie Vychisleniya [Fuzzy Systems and Soft Computing]*, **15**:2 (2020), 83–95(in Russian). https://doi.org/10.26456/fssc72

References

[1] Molodtsov D.A., "Comparison and continuation of multi-valued dependencies", *Nechetkie Sistemy i Myagkie Vychisleniya [Fuzzy Systems and Soft Computing]*, **11**:2 (2016), 115–145 (in Russian).

[2] Molodtsov D.A., "Extrapolation of the multi-valued dependencies", *Nechetkie Sistemy i Myagkie Vychisleniya [Fuzzy Systems and Soft Computing]*, **12**:1 (2017), 45–63 (in Russian).

[3] Molodtsov D.A., "Soft dynamical extrapolation of the multi-valued dependencies", *Nechetkie Sistemy i Myagkie Vychisleniya [Fuzzy Systems and Soft Computing]*, **14**:1 (2019), 5–18 (in Russian).