УДК 004.032.26

Е.В. Бодянский, Б.В. Колчигин, В.В. Волкова, И.П. Плисс

Адаптивная нечеткая кластеризация данных на основе метода Густафсона–Кесселя

Рассмотрены процедуры нечеткой кластеризации. Предложены адаптивные формы вероятностного и возможностного вариантов алгоритма Густафсона–Кесселя, отличающиеся численной простотой и обеспечивающие высокую эффективность при работе в условиях неопределенности, в частности при меняющемся со временем характере обрабатываемых данных и недостатке знаний о природе выборки.

Procedures of fuzzy clustering are considered and adaptive forms of possibilistic and probabilistic Gustafson–Kessel algorithms are suggested. The suggested algorithms are characterized by the numerical simplicity and provide a high efficiency when working in the conditions of uncertainty, particularly in a changing over time character of the data and a lack of knowledge about the nature of a sample.

Розглянуто процедури нечіткої кластеризації. Запропоновано адаптивні форми вирогідністного та можливістного варіантів алгоритму Густафсона–Кесселя, які відрізняються розрахунковою простотою та забезпечують високу ефективність при роботі в умовах невизначеності, а саме при мінливому у часі характері даних, що обробляються, та при нестачі знань про природу вибірки.

Введение. Индуктивное моделирование, по определению А.Г. Ивахненко, – это подбор модели оптимальной сложности для описания имеющихся экспериментальных данных. Сегментация выборки наблюдений на однородные в некотором смысле группы и упрощение их описания путем выделения типичных образцов – необходимый этап решения данной задачи. Этот процесс называется кластеризацией и подразумевает разбиение выборки наблюдений на группы (кластеры), наблюдения внутри которых схожи по некоторому выбранному критерию, а сами группы различны между собой. Традиционный подход к решению этой задачи предполагает, что вся выборка задана заранее, количество кластеров априори известно, а каждое наблюдение может относиться только к одному кластеру.

Ситуация, когда обрабатываемый вектор признаков с различными уровнями принадлежности может относиться сразу к нескольким группам, – предмет рассмотрения нечеткого (фаззи-) кластерного анализа [1–4], в основе которого лежит предположение, что каждому наблюдению может быть приписан некоторый уровень принадлежности каждому кластеру, лежащий в интервале от нуля до единицы [2].

Наиболее общим подходом к решению задачи нечеткой кластеризации есть рассмотрение ее как задачи оптимизации [1, 4], при этом исходная информация – выборка наблюдений, сформированная из N n-мерных векторов признаков $X = \{x_j(k)\}$, $x(k) = (x_1(k), x_2(k), \ldots, x_n(k))^T$.

Результатом кластеризации является $(N \times m)$-матрица нечеткого разбиения $U = \{u_j(k)\}$, задающая уровень принадлежности $u_j(k)$ k-го вектора исходного массива данных $x(k)$ к каждому из m возможных классов ($j = 1, 2, \ldots, m$).

Процедура кластеризации подразумевает расчет центроидов кластеров c_j, обеспечивающих минимизацию некоторого функционала качества разбиения E, в простейшем случае заданного в виде [1]

$$E(u_j, c_j) = \sum_{k=1}^{N} \sum_{j=1}^{m} u_j^{\beta}(k) d^2\left(x(k), c_j\right), \qquad (1)$$

при ограничении

$$\sum_{j=1}^{m} u_j(k) = 1, \qquad (2)$$

где β – весовой коэффициент, определяющий степень размытия кластеров (фаззификатор) и принимающий значения из интервала $[1, \infty)$, а функция $d^2(x(k), c_j)$ определяет расстояние ме-

жду k-м наблюдением и центром j-го кластера в принятой метрике.

Базовые алгоритмы нечеткой кластеризации

Решение задачи нечеткой кластеризации в описанной постановке можно свести к оптимизации функции Лагранжа

$$L\left(u_j(k), c_j, \lambda(k)\right) = \\ = \sum_{k=1}^{N}\sum_{j=1}^{m} u_j^{\beta}(k)\, d^2\left(x(k), c_j\right) + \\ + \sum_{k=1}^{N}\lambda(k)\left(\sum_{j=1}^{m} u_j(k) - 1\right), \tag{3}$$

где $\lambda(k), k = 1, 2, \ldots, N$ – неопределенные множители Лагранжа. Важно обратить внимание на то, что выбранная функция расчета расстояния (метрика) $d^2\left(x(k), c_j\right)$ явно входит в оптимизируемую функцию (3), а значит ее изменение непосредственно в алгоритме невозможно: изменение метрики порождает новый алгоритм.

В случае использования квадрата евклидовой метрики, т.е. при $d^2\left(x(k), c_j\right) = \left(x(k) - c_j\right)^T \times \\ \times \left(x(k) - c_j\right) = \left\| x(k) - c_j \right\|^2$, решение системы уравнений Каруша–Куна–Таккера ведет к известному результату

$$\begin{cases} u_j(k) = \dfrac{\left(\left\| x(k) - c_j \right\|^2\right)^{\frac{1}{1-\beta}}}{\sum_{l=1}^{m}\left(\left\| x(k) - c_l \right\|^2\right)^{\frac{1}{1-\beta}}}, \\ c_j = \dfrac{\sum_{k=1}^{N} u_j^{\beta}(k)\, x(k)}{\sum_{k=1}^{N} u_j^{\beta}(k)}, \\ \lambda(k) = -\left(\left(\sum_{l=1}^{m}\beta \left\| x(k) - c_l \right\|^2\right)^{\frac{1}{1-\beta}}\right)^{1-\beta}, \end{cases} \tag{4}$$

что при $\beta = 2$ дает алгоритм нечетких C-средних Бездека [1] (*fuzzy C-means*, *FCM*):

$$\begin{cases} u_j(k) = \dfrac{\left\| x(k) - c_j \right\|^{-2}}{\sum_{l=1}^{m}\left\| x(k) - c_l \right\|^{-2}}, \\ c_j = \dfrac{\sum_{k=1}^{N} u_j^2(k)\, x(k)}{\sum_{k=1}^{N} u_j^2(k)}. \end{cases} \tag{5}$$

Алгоритм нечетких C-средних – самый популярный из алгоритмов нечеткой кластеризации, однако имеет некоторые ограничения и сильные априорные допущения о характере данных в обрабатываемой выборке, сужающие область его применимости.

Одно из ограничений алгоритма нечетких C-средних происходит от использования функции принадлежности, определяемой евклидовой метрикой. Несложно заметить, что значение принадлежности $u_j(k)$, вычисленное по алгоритму (5), будет одинаковым для всех точек, находящихся на одинаковом расстоянии от центроида c_j, а это значит, что формируемые кластеры будут иметь строго сферическую форму. В то же время в реальных задачах нередко встречаются данные, в которых форма скоплений сильно отличается от (гипер) сферы, либо шкалы значений отдельных компонент векторов наблюдений имеют разный масштаб. Априорное предположение о сферической форме кластеров делает в таких случаях глобальный минимум функционала (1) недостижим для алгоритма нечетких C-средних.

Обобщением алгоритма нечетких C-средних есть алгоритм Густафсона–Кесселя [5], способный выделять кластеры эллипсоидальной формы. Для этого в формулу расчета расстояний между векторами вводится масштабирующая матрица A, меняющая масштаб пространства по каждой координате:

$$d_A^2(x, c) = (x - c)^T A (x - c) = \left\| x - c \right\|_A^2. \tag{6}$$

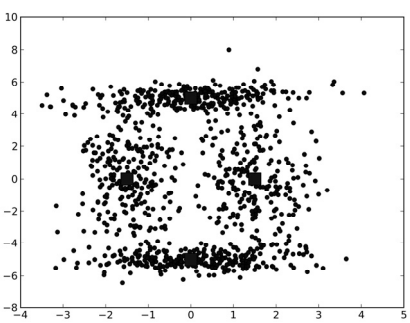
Используя в функции Лагранжа (3) метрику (6), получаем известную процедуру минимизации, включающую, помимо расчета центров кластеров и уровней принадлежности, также вычисление ковариационной матрицы каждого

кластера S_j и, на основании ее, масштабирующей матрицы кластера A_j в виде:

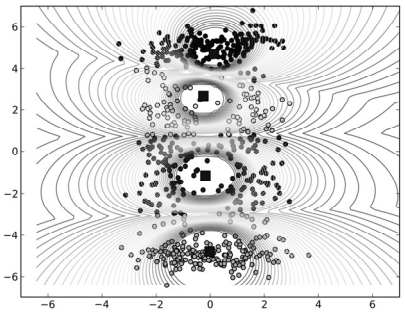
$$\begin{cases} u_j(k) = \dfrac{\left(\left(x(k)-c_j(k)\right)^T A_j \left(x(k)-c_j(k)\right)\right)^{\frac{1}{1-\beta}}}{\sum\limits_{l=1}^{m}\left(\left(x(k)-c_l(k)\right)^T A_l \left(x(k)-c_j(k)\right)\right)^{\frac{1}{1-\beta}}} = \dfrac{\left\|x(k)-c_j\right\|_{A_j}^{\frac{2}{1-\beta}}}{\sum\limits_{l=1}^{m}\left\|x(k)-c_j\right\|_{A_j}^{\frac{2}{1-\beta}}}, \\ c_j = \dfrac{\sum\limits_{k=1}^{N} u_j^{\beta}(k)x(k)}{\sum\limits_{k=1}^{N} u_j^{\beta}(k)}, \\ S_j = \sum\limits_{k=1}^{N} u_j^{\beta}(k)\left(x(k)-c_j\right)\left(x(k)-c_j\right)^T, \\ A_j = |S_j|^{\frac{1}{n}} S_j^{-1}. \end{cases} \quad (7)$$

Полученная таким образом матрица A_j является симметричной и положительно определенной, а задаваемый ею масштаб, определяемый для каждой координаты квадратным корнем соответствующего собственного вектора, – оптимальным [6] с учетом критерия (1). Значения собственных векторов матрицы A_j способны изменять форму соответствующего кластера, что значительно расширяет область применимости алгоритма.

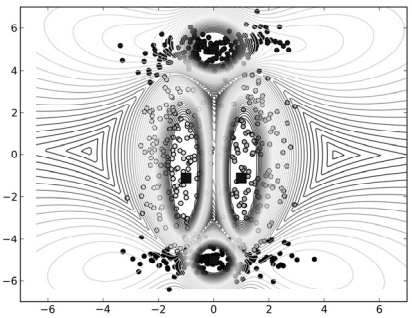
Иллюстрация разницы работы алгоритмов показана на рис. 1. Для этого создана искусственная выборка из 1000 точек, бинормально распределенных в окрестностях четырех центров (рис. 1,*а*). Особенностью данной выборки есть то, что граничные данные вытянутых кластеров лежат ближе к центру соседнего кластера, чем своего собственного. Поэтому применение алгоритма нечетких *C*-средних не приводит к ожидаемому разбиению (рис. 1,*б*). В то же время заметно, что центры и параметры распределения (ковариационные матрицы), воссозданные алгоритмом Густафсона–Кесселя (рис. 1,*в*) предельно близки исходно заданным.



а – исходная выборка



б – алгоритм FCM



в– алгоритм Густафсона–Кесселя

Рис. 1. Сравнение результатов работы нечетких алгоритмов кластеризации на модельной выборке

Возможностные алгоритмы нечеткой кластеризации

Одним из основных недостатков алгоритма нечетких *C*-средних есть его низкая устойчивость к выбросам в данных, а также к существенно перекрывающимся кластерам. Эти недостатки следуют из ограничения, которое связано с интерпретацией степени принадлежности как значения вероятности того, что наблюдение принадлежит к кластеру (поэтому алгоритмы, использующие это ограничение, называют *вероятностными*). Очевидно, что условие (2) слишком жесткое для многих реальных задач: выбросы в данных должны иметь близкую к нулю степень принадлежности ко всем кластерам, а наблюдениям, находящимся в центре существенно перекрывающихся групп, уместно было бы присвоить принадлежность, близкую к единице, к каждой из них. И то, и другое запрещено ограничением (2), требующим равенства единице суммы принадлежностей отдельного наблюдения всем кластерам. Алгоритмы, не использующие это ограничение, называются *возможностными* (*possibilistic*) нечеткими алгоритмами [7, 8].

В возможностных алгоритмах кластеризации базовая целевая функция имеет вид

$$E\left(u_j(k),c_j\right)=\sum_{k=1}^{N}\sum_{j=1}^{m}u_j^{\beta}\left(k\right)d^2\left(x(k),c_j\right)+$$

$$+\sum_{j=1}^{m}\mu_j\sum_{k=1}^{N}(1-u_j\left(k\right))^{\beta}, \qquad (8)$$

где параметр μ_j задает нечеткую границу j-го кластера, т.е. если $d^2\left(x(k),c_j\right)=\mu_j$, то $u_j(k)=0{,}5$. При использовании евклидовой метрики, получаем минимизирующую процедуру

$$\begin{cases} u_j\left(k\right)=\left(1+\left(\dfrac{\left\|x\left(k\right)-c_j\right\|^2}{\mu_j}\right)^{\frac{1}{\beta-1}}\right)^{-1}, \\ c_j=\dfrac{\sum\limits_{k=1}^{N}u_j^{\beta}\left(k\right)x(k)}{\sum\limits_{k=1}^{N}u_j^{\beta}\left(k\right)}, \\ \mu_j\left(k\right)=\dfrac{\sum\limits_{k=1}^{N}u_j^{\beta}\left(k\right)\left\|x\left(k\right)-c_j\right\|^2}{\sum\limits_{k=1}^{N}u_j^{\beta}\left(k\right)}, \end{cases} \qquad (9)$$

известную при $\beta=2$ как возможностный алгоритм C-средних [7] (*possibilistic C-means, PCM*).

Из-за большего количества рассчитываемых параметров алгоритм Густафсона–Кесселя еще более неустойчив к выбросам и не менее алгоритма нечетких C-средних уязвим в ситуации существенно перекрывающихся кластеров. Аналогично алгоритму нечетких C-средних, мы можем уйти от этих недостатков, перейдя к возможностному алгоритму Густафсона–Кесселя [7]. Для этого достаточно заменить метрику в критерии (8) на масштабируемую метрику (6), что приводит к известной минимизирующей процедуре:

$$\begin{cases} u_j(k)=\left(1+\left(\dfrac{(x(k)-c_j)^T A_j(x(k)-c_j)}{\mu_j}\right)^{\frac{1}{1-\beta}}\right)^{-1}= \\ \quad=\left(1+\left(\dfrac{\left\|x(k)-c_j\right\|_{A_j}^2}{\mu_j}\right)^{\frac{1}{1-\beta}}\right)^{-1}, \\ c_j=\dfrac{\sum\limits_{k=1}^{N}u_j^{\beta}(k)x(k)}{\sum\limits_{k=1}^{N}u_j^{\beta}(k)}, \\ S_j=\sum\limits_{k=1}^{N}u_j^{\beta}(k)\left(x(k)-c_j\right)\left(x(k)-c_j\right)^T \\ A_j=|S_j|^{\frac{1}{n}}S_j^{-1}, \\ \mu_j(k)=\dfrac{\sum\limits_{k=1}^{N}u_j^{\beta}(k)(x(k)-c_j)^T A_j(x(k)-c_j)}{\sum\limits_{k=1}^{N}u_j^{\beta}(k)}= \\ \quad=\dfrac{\sum\limits_{k=1}^{N}u_j^{\beta}(k)\left\|x(k)-c_j\right\|_{A_j}^2}{\sum\limits_{k=1}^{N}u_j^{\beta}(k)}. \end{cases} \qquad (10)$$

Важно отметить, что у возможностных алгоритмов есть свои недостатки. В первую очередь в отсутствии ограничения (2) ничто не мешает центроидам сближаться или отдаляться неограниченно. Вследствие этого результат таких алгоритмов может отличаться нестабильностью, т.е. малые возмущения в исходных данных (в первую очередь в инициализации центроидов) ведут к значительной разнице в результирующих разбиениях. Основной рекомендацией для борьбы с этим обстоятельством есть использование для инициализации возможностного алгоритма результата работы нескольких итераций алгоритма вероятностного.

Адаптивные алгоритмы нечеткой кластеризации

Вышеописанные процедуры реализуют пакетную обработку данных, т.е. подразумевают наличие заданной и не меняющейся в процессе работы выборки из N наблюдений, а также заданное заранее количество кластеров. Это делает их неприменимыми в задачах, когда дан-

ные поступают на обработку последовательно в *on-line* режиме. Для работы в этих условиях в [9–12] предложены адаптивные (рекуррентные) формы некоторых алгоритмов кластеризации. В частности, применяя к (3) процедуру нелинейного программирования Эрроу–Гурвица–Удзавы, можно получить рекуррентный алгоритм нечеткой кластеризации по критерию (1) с евклидовой метрикой [12]:

$$\begin{cases} u_j(k+1) = \dfrac{\left(\left\|x(k+1)-c_j(k)\right\|\right)^{\frac{2}{1-\beta}}}{\sum_{l=1}^{m}\left(\left\|x(k+1)-c_l(k)\right\|\right)^{\frac{2}{1-\beta}}}, \\ c_j(k+1) = c_j(k) + \eta(k)u_j^{\beta}(k+1)\left(x(k+1)-c_j(k)\right), \end{cases} \quad (11)$$

являющийся обобщением алгоритма обучения Чанга–Ли [10, 11] и совпадающий при $\beta = 2$ с градиентной процедурой кластеризации Парка–Дэггера [9]. Можно отметить, что при $\beta = 1$ процедура (11) совпадает с четким алгоритмом *K*-средних (*K-means*).

Таким образом, функционирующая по такому алгоритму система воспринимает на входе лишь одно наблюдение $x(k+1)$, рассчитывает только его принадлежность к *j*-му кластеру $u_j(k+1)$ и сразу же вычисляет новое положение соответствующего центроида $c_j(k+1)$.

Используя в качестве метрики (6), аналогичным образом можно получить рекуррентную форму алгоритма Густафсона–Кесселя в виде:

$$\begin{cases} u_j(k+1) = \dfrac{\left\|x(k+1)-c_j(k)\right\|_{A_j(k)}^{\frac{2}{1-\beta}}}{\sum_{l=1}^{m}\left\|x(k+1)-c_l(k)\right\|_{A_j(k)}^{\frac{2}{1-\beta}}}, \\ S_j(k+1) = S_j(k) + u_j^{\beta}(k+1)\left(x(k+1)-c_j(k)\right)\left(x(k+1)-c_j(k)\right)^T, \\ S_j^{-1}(k+1) = S_j^{-1}(k) - \\ \quad - \dfrac{u_j^{\beta}(k+1)S_j^{-1}(k)\left(x(k+1)-c_j(k)\right)\left(x(k+1)-c_j(k)\right)^T S_j^{-1}(k)}{1+u_j^{\beta}(k+1)\left(x(k+1)-c_j(k)\right)^T S_j^{-1}(k)\left(x(k+1)-c_j(k)\right)}, \\ \left|S_j(k+1)\right| = \left|S_j(k)\right|\left(1+u_j^{\beta}(k+1)\left\|x(k+1)-c_j(k)\right\|^2\right), \\ A_j(k+1) = \left|S_j(k+1)\right|^{\frac{1}{n}} S_j^{-1}(k+1), \\ c_j(k+1) = c_j(k) + \eta(k)u_j^{\beta}(k+1)A_j(k+1)\left(x(k+1)-c_j(k)\right). \end{cases} \quad (12)$$

Особенность данного алгоритма – рекуррентное обращение ковариационной матрицы $S_j^{-1}(k+1)$ по формуле Шермана–Моррисона и рекуррентное вычисление ее определителя $\left|S_j(k+1)\right|$.

Адаптивной форме вероятностных алгоритмов, несмотря на удобство использования, присущи те же недостатки, что и их пакетной форме – уязвимость к выбросам и сложности с разделением сильно пересекающихся кластеров. Кроме того, такие алгоритмы не могут быть названы в полной мере алгоритмами обучения без учителя, поскольку экспериментатор все еще должен заранее задать количество кластеров, используя знания о природе выборки. Для компенсации этих недостатков логично обратиться к адаптивной форме возможностных алгоритмов нечеткой кластеризации.

В [12] описан синтез адаптивного возможностного алгоритма нечетких *C*-средних с применением процедуры нелинейного программирования Эрроу–Гурвица–Удзавы к критерию (8) при использовании евклидовой метрики. Этот алгоритм имеет простой вид:

$$\begin{cases} u_j(k) = \left(1+\left(\dfrac{\left\|x(k)-c_j\right\|^2}{\mu_j(k)}\right)^{\frac{1}{1-\beta}}\right)^{-1}, \\ c_j(k+1) = c_j(k) + \eta(k)u_j^{\beta}(k+1)(x(k+1)-c_j(k)), \\ \mu_j(k+1) = \dfrac{\sum_{p=1}^{k+1} u_j^{\beta}(p)\left\|x(p)-c_j(k+1)\right\|^2}{\sum_{p=1}^{k+1} u_j^{\beta}(p)}. \end{cases} \quad (13)$$

Благодаря отсутствию ограничения (2), возможно не только отслеживать выбросы, но и немедленно реагировать на них, например, создавая новый центроид, если сумма принадлежностей одного или нескольких наблюдений к существующим кластерам слишком мала. Это и позволяет назвать алгоритм *адаптивным*.

Логичным шагом будет использование преимуществ адаптивного и возможностного под-

ходов к алгоритму Густафсона–Кесселя. Использование в качестве метрики (6) приводит к следующей процедуре:

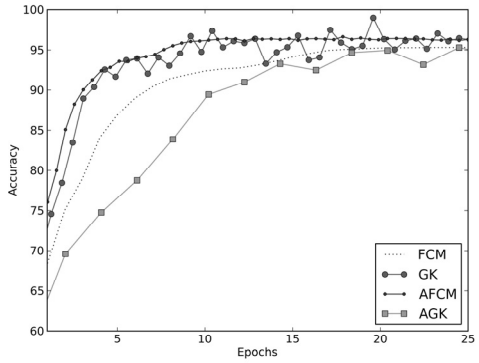
$$
\begin{cases}
u_j(k+1) = \left(1 + \left(\dfrac{\left\| x(k+1) - c_j(k) \right\|^2_{A_j(k)}}{\mu_j(k)} \right)^{\frac{1}{1-\beta}} \right)^{-1}, \\
S_j(k+1) = S_j(k) + u_j^{\beta}(k+1)\left(x(k+1) - c_j(k) \right)\left(x(k+1) - c_j(k) \right)^T, \\
S_j^{-1}(k+1) = S_j^{-1}(k) - \\
\quad - \dfrac{u_j^{\beta}(k+1) S_j^{-1}(k)\left(x(k+1) - c_j(k) \right)\left(x(k+1) - c_j(k) \right)^T S_j^{-1}(k)}{1 + u_j^{\beta}(k+1)\left(x(k+1) - c_j(k) \right)^T S_j^{-1}(k)\left(x(k+1) - c_j(k) \right)}, \\
\left| S_j(k+1) \right| = \left| S_j(k) \right| \left(1 + u_j^{\beta}(k+1) \left\| x(k+1) - c_j(k) \right\|^2 \right), \\
A_j(k+1) = \sqrt[n]{\left| S_j(k+1) \right|} S_j^{-1}(k+1), \\
c_j(k+1) = c_j(k) + \eta(k) u_j^{\beta}(k+1) A_j(k+1)\left(x(k+1) - c_j(k) \right), \\
\mu_j(k+1) = \dfrac{\sum\limits_{p=1}^{k+1} u_j^{\beta}(p) \left\| x(k+1) - c_j(k) \right\|^2_{A_j(k+1)}}{\sum\limits_{p=1}^{k+1} u_j^{\beta}(p)}.
\end{cases}
\tag{14}
$$

Видно, что за исключением вида функции принадлежности $u_j(k+1)$ и вычисления связанного с ней параметра нечеткой границы $\mu_j(k+1)$, алгоритм близок к адаптивному вероятностному алгоритму Густафсона–Кесселя (12).

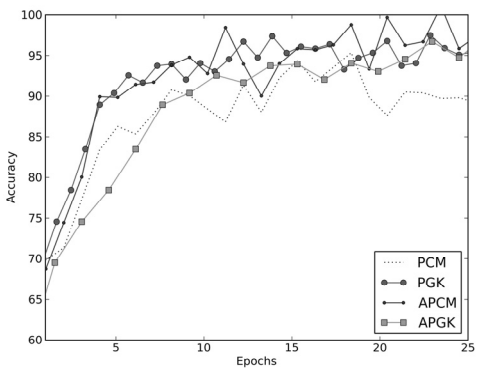
Экспериментальные исследования

Ввиду того что точность и иные меры качества кластеризации адаптивных алгоритмов идентичны по построению их пакетным аналогам, наиболее интересной для экспериментального исследования была признана скорость самообучения системы. В качестве меры модельного времени, как минимальный общий квант пакетных и адаптивных форм алгоритмов, принято количество проходов (эпох, итераций) по всей доступной выборке наблюдений. В проведенной серии экспериментов тестировалось время, за которое система достигает заданной точности кластеризации. Для тестирования использовалась традиционная выборка «Вина» (178 13-мерных наблюдений, разделенных на три класса).

С каждым из рассмотренных алгоритмов была проведена серия из 50 экспериментов. Каждый эксперимент включал 25 итераций обучения алгоритма. Изначально инициализированный случайно, на каждой итерации алгоритм самообучался на обучающем множестве, включающем 66 процентов выборки, после чего измерялась точность кластеризации на всей выборке. На графиках рис. 2 показана средняя точность кластеризации каждого алгоритма в зависимости от количества проходов по выборке.



а – вероятностные алгоритмы



б – возможностные алгоритмы

Рис. 2. Точность кластеризации в зависимости от числа итераций по выборке. Аббревиатурами обозначены алгоритмы: *FCM* – нечетких *C*-средних, *GK* – Густафсона–Кесселя, *PCM* – возможностный нечетких *C*-средних, *PGK* – возможностный Густафсона–Кесселя. Буква «А» обозначает адаптивные варианты соответствующих алгоритмов

Отметим, что из-за большого количества вычисляемых параметров, адаптивные формы как возможностного, так и вероятностного алгоритма Густафсона–Кесселя, требуют для качественной настройки даже большего количества наблюдений, чем их пакетные формы, что обычно нехарактерно для адаптивных алгоритмов, однако закономерно с учетом сложности систем. Значительно важнее, что, имея намного более гибкие возможности адаптации к входящим наблюдениям, адаптивные формы алгоритмов Густафсона–Кесселя сохраняют монотонность

роста качества кластеризации с числом принятых наблюдений даже лучше большинства пакетных алгоритмов, что особенно выгодно выделяет адаптивный возможностный алгоритм Густафсона–Кесселя на фоне других возможностных алгоритмов, традиционно отличающихся нестабильностью результатов.

Заключение. Предложены адаптивные процедуры вероятностного и возможностного вариантов алгоритма нечеткой кластеризации Густафсона–Кесселя, отличающиеся простотой численной реализации и возможностью обработки данных в режиме реального времени. Экспериментальные данные подтверждают работоспособность предложенных алгоритмов и уникальность некоторых их свойств: не требуют ручной настройки, способны обрабатывать выбросы без предобработки данных и адаптироваться к меняющимся характеристикам выборки. Дальнейшие исследования могут быть направлены на автоматическую настройку параметров фаззификатора и шага обучения.

1. *Bezdek J.C*. Pattern Recognition with Fuzzy Objective Function Algorithms. – N.Y.: Plenum Press. – 1981. – 272 p.
2. *Fuzzy* Clustering Analysis: Methods for Classification, Data Analysis and Image Recognition / F. Höppner, F. Klawonn, R. Kruse et al. – Chichester: John Willey & Sons. – 1999. – 289 p.
3. *Gan G*., *Ma Ch*., *Wu J*. Data Clustering: Theory, Algorithms and Applications – Philadelphia: SIAM, 2007. – 466 p.
4. *Gath I*., *Geva A.B*. Unsupervised optimal fuzzy clustering // Patt. Analysis and Machine Intell. – 1989. – **11**, 7. – P. 773–780.
5. *Gustafson D.E*., *Kessel W.C*. Fuzzy clustering with a fuzzy covariance matrix. // IEEE Conf. on Decision and Control including the 17th Symp. on Adaptive Proc. – 1978. – P. 761–766.
6. *Krishnapuram R*., *Jongwoo K*. A note on the Gustafson–Kessel and adaptive fuzzy clustering algorithms // IEEE Trans. on Fuzzy Syst. – **7**, N 4. – Aug. 1999. – P. 453–461.
7. *Krishnapuram R*., *Keller J.M*. A possibilistic approach to clustering // Fuzzy Systems. – 1993. – **1**, N 2. – P. 98–110.
8. *Krishnapuram R*., *Keller J.M*. Fuzzy and possibilistic clustering methods for computer vision // Neural Fuzzy Syst. – 1994. – **12**. – P. 133–159.
9. *Park D.C*., *Dagher I*. Gradient based fuzzy c-means (GBFCM) algorithm // Proc. IEEE Int. Conf. on Neural Networks. – 1984. – P. 1626–1631.
10. *Chung F.L*., *Lee T*. Fuzzy competitive learning // Neural Networks. – 1994. – **7**, N 3. – P. 539–552.
11. *Chung F.-L*., *Lee T*. Unsupervised fuzzy competitive learning with monotonically decreasing fuzziness // Proc. 1993 Int. Joint Conf. on Neural Networks. – 1993. – P. 2929–2932.
12. *Bodyanskiy Ye*. Computational intelligence techniques for data analysis / Lecture Notes in Informatics. – Bonn, Germany. – 2005. – **P–72**. – P. 15–36.

Тел. для справок: +38 067 399-9150, +38 063 319-6020, +38 057 702-1337, 702-1890 (Харьков)
E-mail: *bodya@kture.kharkov.ua*, *quasimail@gmail.com*, *volkova@kture.kharkov.ua*, *pliss@kture.kharkov.ua*

© Е.В. Бодянский, Б.В. Колчигин, В.В. Волкова, И.П. Плисс, 2013