

3. Кравченко Р.Г. Математическое моделирование экономических процессов в сельском хозяйстве. – М. : Колос, 1978. – 424 с.

4. Математическое моделирование экономических процессов в сельском хозяйстве / Гатаулин А.М., Гаврилов Г.В., Сорокин Т.М. и др. ; под ред. А.М.Гатаулина. – М. : Агропромиздат, 1990. – 432 с.

Об улучшении алгоритма k -средних

В.А. Ольховатских

БГПУ, г. Барнаул

Алгоритм k -средних относится к итеративным методам группировки кластерного анализа. Идея алгоритма состоит в минимизации суммарного отклонения по каждому кластеру. Алгоритм k -средних разработал J.B. MacQueen в 1967 году (см [1]). Полное описание алгоритма можно найти в работе [2].

В качестве объектов наблюдений будем рассматривать точки на евклидовой плоскости с евклидовой метрикой.

Стандартный алгоритм k -средних:

1. Случайно выбрать k начальных «центров масс» кластеров (любые k из n объектов, или вообще k случайных точек).
2. Отнести каждый объект к кластеру с ближайшим «центром масс».
3. Пересчитать «центры масс» кластеров согласно текущему членству.
4. Если критерий останова алгоритма не удовлетворен, вернуться к шагу 2.

(Критерий останова – стабилизация процесса, т.е. неизменность центров тяжести кластеров).

В связи с произвольностью выбора начальных «центров масс», результаты кластеризации одного и того же множества могут сильно отличаться друг от друга.

Для улучшения алгоритма предлагается в качестве начальных «центров масс» выбирать точки получаемые следующим образом:

1. Каждый объект считается отдельным кластером с весом равным 1.
2. Находятся два самых близких объекта p_i и p_j ($i \neq j$).
3. Находится «центр масс» c объектов p_i и p_j по формуле

$$c = \frac{m_i p_i + m_j p_j}{m_i + m_j}, \text{ где } m_i \text{ и } m_j \text{ веса } p_i \text{ и } p_j \text{ соответственно.}$$

4. Элементы p_i и p_j исключаются из рассмотрения.

5. «Центр масс» c добавляется в рассмотрение как объект множества с массой равной $m_i + m_j$.
6. Если критерий остановки алгоритма не удовлетворен, вернуться к шагу 2.

(Критерий остановки – количество нерассмотренных элементов равно k – количеству кластеров).

В итоге получим k точек, которые выбираем в качестве начальных «центров масс» кластеров. Вследствие применения данного алгоритма результаты кластеризации одного и того же множества будут одинаковыми, а общая внутриклассовая дисперсия (ОВД) уменьшится.

Эмпирически получено, что ОВД данного алгоритма на 0,5%–5% меньше ОВД стандартного алгоритма и на 0,2%–3% меньше алгоритма математического пакета Matlab v7.6.0 R2008a.

Литература

1. MacQueen J.B. «Some Methods for classification and Analysis of Multivariate Observations», Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1967. 1:281-297
2. Hartigan J. A. Clustering algorithms. N.Y., 1975.
3. Мандель И. Д. Кластерный анализ. – М.: Финансы и статистика. 1988.–176 с.:ил.
4. Бородич С. А. Вводный курс эконометрики : учебное пособие – Мн.: БГУ, 2000. – 354 с.
5. Факторный, дискриминантный и кластерный анализ : пер. с англ. / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др.; под ред. И. С. Енюкова. – М. : Финансы и статистика, 1989.– 215 с. : ил.

Моделирование равновесного состояния на товарных рынках

А.Б. Панюкова
АлтГУ, г. Барнаул

Исследование взаимодействия экономических агентов на товарном рынке с торговыми посредниками в условиях рыночной экономики является одной из важнейших задач повышения эффективности функционирования предпринимательских структур, поскольку в большин-