

**КОВАЛЕНКО А. П.**

**СОСТОЯТЕЛЬНАЯ ОЦЕНКА КЛАСТЕРОВ  
ВЫСОКОЙ ПЛОТНОСТИ\***

В статье приводится определение кластеров высокой плотности относительно функции уровня. Предлагается состоятельный алгоритм построения таких кластеров по выборке, основанный на непараметрических оценках плотности распределения. Приводятся результаты решения численного примера.

**1. Введение.** Кластерный анализ включает в себя совокупность моделей, методов и алгоритмов, предназначенных для разбиения совокупности объектов на некоторое число классов так, чтобы в один класс попали «похожие» в определенном смысле объекты. Исходная информация об объектах обычно представлена в форме матрицы «объект–признак», каждая строка  $x_i$  которой есть результат статистического обследования реального объекта по  $d$  анализируемым переменным, либо в форме матрицы попарных расстояний между объектами. Если исходные данные  $x_i, i = 1, \dots, n$ , интерпретируются как наблюдения многомерных случайных векторов, то говорят о теоретико-вероятностной постановке задачи кластерного анализа (см., например, [1]). При этом ключевым является определение понятия теоретического кластера. Подробный обзор теоретико-вероятностных моделей кластерного анализа приведен в [2]. Ниже рассматривается одна из таких моделей – модель кластеров высокой плотности.