3

Objective Function Clustering

S8 illustrates some of the difficulties inherent with cluster analysis; its aim is to alert investigators to the fact that various algorithms can suggest radically different substructures in the same data set. The balance of Chapter 3 concerns objective functional methods based on fuzzy c-partitions of finite data. The nucleus for all these methods is optimization of nonlinear objectives involving the weights $\{u_{ik}\}$; functionals using these weights will be differentiable over M_{fc}—but not over M_c—a decided advantage for the fuzzy embedding of hard c-partition space. Classical first- and second-order conditions yield iterative algorithms for finding the optimal fuzzy c-partitions defined by various clustering criteria.

S9 describes one of Ruspini's[(89)] algorithms, and compares its performance to the hard c-means or classical within-groups sum-of-squared-errors method. S10 discusses an algorithm based on a mixed fuzzy statistical objective due to Woodbury and Clive[(118)] and reports an application of it to the diagnosis of tetralogy of Fallot (four variants of a congenital heart disease). The infinite family of fuzzy c-means algorithms is defined in S11 and is compared (numerically) to several earlier methods. S12 contains a proof of convergence for fuzzy c-means; and S13 illustrates its usefulness for feature selection with binary data by examining a numerical example concerning stomach disorders due to hiatal hernia and gallstones.

S8. Cluster Analysis: An Overview

In this chapter, we assume that the important question of feature extraction—which characteristics of the physical process are significant indicators of structural organization, and how to obtain them—has been answered. Our point of departure is this: given $X = \{x_1, x_2, \ldots, x_n\}$, find an integer c, $2 \leq c < n$, and a c-partition of X exhibiting categorically homogeneous subsets! The most important requirement for resolving this issue is a suitable measure of "clusters"—what *clustering criterion* shall be used?

J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*

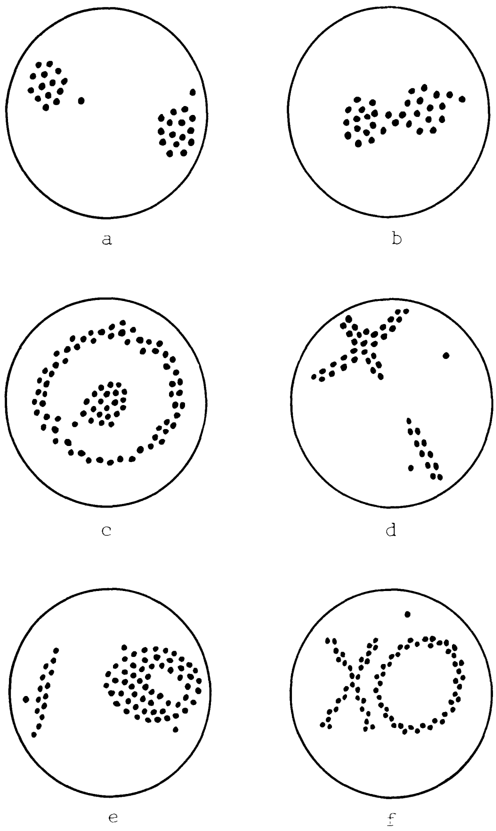

Figure 8.1. Some clustered data sets in the plane.

Specifically, what mathematical properties—e.g., distance, angle, curvature, symmetry, connectivity, intensity—possessed by members of the data should be used, and in what way, to identify the clusters in X? Since each observation can easily have several hundred dimensions, the variety of "structures" is without bound. It is clear that (i) *no* clustering criterion or measure of similarity will be universally applicable, and (ii) selection of a particular criterion is at least partially subjective, and always open to question.

Regardless of the actual dimension of data space, clustering models are usually interpreted geometrically by considering their action on two- or three-dimensional examples. Some of the difficulties inherent in trying to formulate a successful clustering criterion for a wide variety of data structures are illustrated in Fig. 8.1.

The ideal case—compact, well-separated, equally proportioned clusters (Fig. 8.1a)—is seldom encountered in real data. More realistically, data