Many studies have been done with respect to variations and generalizations of the basic methods of fuzzy \(c\)-means. We will divide those variations and generalizations into two classes. The first class has ‘standard variations or generalizations’ that include relatively old studies, or should be known to many readers of general interest. On the other hand, the second class includes more specific studies or those techniques for a limited purpose and will be interested in by more professional readers. We describe some algorithms in the first class in this chapter.

### 3.1 Possibilistic Clustering

Krishnapuram and Keller \[87\] propose the method of possibilistic clustering: the same alternate optimization algorithm \textbf{FCM} is used in which the constraint \(U_f\) is not employed but nontrivial solution of

\[
\sum_{k=1}^{N} u_{ki} > 0, \quad 1 \leq i \leq c; \quad u_{kj} \geq 0, \quad 1 \leq k \leq n, \quad 1 \leq j \leq c
\]  

(3.1)

should be obtained.

For this purpose the objective function \(J_{\text{fcm}}\) cannot be used since the optimal \(\bar{U}\) is trivial: \(\bar{u}_{ki} = 0\) for all \(i\) and \(k\). Hence a modified objective function

\[
J_{\text{pos}}(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{N} (u_{ki})^m D(x_k, v_i) + \sum_{i=1}^{c} \eta_i \sum_{k=1}^{N} (1 - u_{ki})^m
\]

(3.2)

has been proposed. The solution \(\bar{U}\) becomes

\[
\bar{u}_{ki} = \frac{1}{1 + \left( \frac{D(x_k, \bar{v}_i)}{\eta_i} \right)^m}
\]

(3.3)
while the optimal $\bar{v}_i$ remains the same:

$$\bar{v}_i = \frac{\sum_{k=1}^{N} (\bar{u}_{ki})^m x_k}{\sum_{k=1}^{N} (\bar{u}_{ki})^m}.$$

Notice that the fuzzy classification function derived from the above $\bar{U}$ is

$$U_{\text{pos}}(x; v_i) = \frac{1}{1 + \left( \frac{D(x, v_i)}{\eta_i} \right)^{\frac{m}{m-1}}} \quad (3.4)$$

We will observe other types of possibilistic clustering in which we obtain different classification functions. A classification function from a method of possibilistic clustering in general is denoted by $U(x; v_i)$. Notice that this form is different from that for fuzzy $c$-means: the latter is $U_{(i)}(x; V)$ with the superscript $(i)$ and the parameter $V$, while the former is without the superscript and the parameter is just $v_i$.

The classification function $U_{\text{pos}}(x; v_i)$ has the next properties, when we put $U(x; v_i) = U_{\text{pos}}(x; v_i)$.

(i) $U(x; v_i)$ is unimodal with the maximum value at $x = v_i$.

(ii) $\max_{x \in R^p} U(x; v_i) = U(v_i; v_i) = 1$.

(iii) $\inf_{x \in R^p} U(x; v_i) = 0$.

(iv) Let us define the set $X_i \subset R^p$ by

$$X_i = \{ x \in R^p : U(x; v_i) > U(x; v_j), \forall j \neq i \}.$$

Then, $X_1, \ldots, X_c$ are the Voronoi sets of $R^p$.

### 3.1.1 Entropy-Based Possibilistic Clustering

The above objective function should not be the only one for the possibilistic clustering [23]; the entropy-based objective function (6.4) can also be used for the possibilistic method:

$$J_{efc}(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ki} D(x_k, v_i) + \nu \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ki} \log u_{ki}.$$

The solution $\bar{U}$ and $\bar{V}$ are respectively given by

$$\bar{u}_{ki} = \exp \left( -1 - \frac{D(x_k, \bar{v}_i)}{\nu} \right) \quad (3.5)$$