4.2 Convex Sets, Convex Hulls and Linear Separability

Consider two pattern sets $\mathcal{X}_1$ and $\mathcal{X}_2$ sampled from two classes $\mathcal{C}_1$ and $\mathcal{C}_2$. Many subsets of $\mathbb{R}^n$ contain the pattern sets $\mathcal{X}_1$ and $\mathcal{X}_2$. In the present context we are interested only in the smallest convex sets that contain these sets, and very shortly we will see why. But first the following definitions are in order.

Recall that a vector of input features is called a *pattern*. The vector space from which valid patterns may be derived is then the *pattern space*, typically $\mathbb{R}^n$. In general any n–dimensional pattern can be represented by a point in pattern space, $\mathbb{R}^n$.

Definition 4.2.1

Let $X, Y \in \mathcal{S} \subset \mathbb{R}^n$, then $\mathcal{S}$ is convex iff $\lambda X + (1-\lambda)Y \in \mathcal{S}$, $0 \le \lambda \le 1$, $\forall X, Y \in \mathcal{S}$. Equivalently, a set $\mathcal{S}$ is convex if it contains all points on all line segments with end points in $\mathcal{S}$.

In $\mathbb{R}^n$, spheres, ellipsoids and cubes are examples of convex sets. Tori and disconnected sets are not convex. Figure 4.1 portrays some convex and non-convex sets.

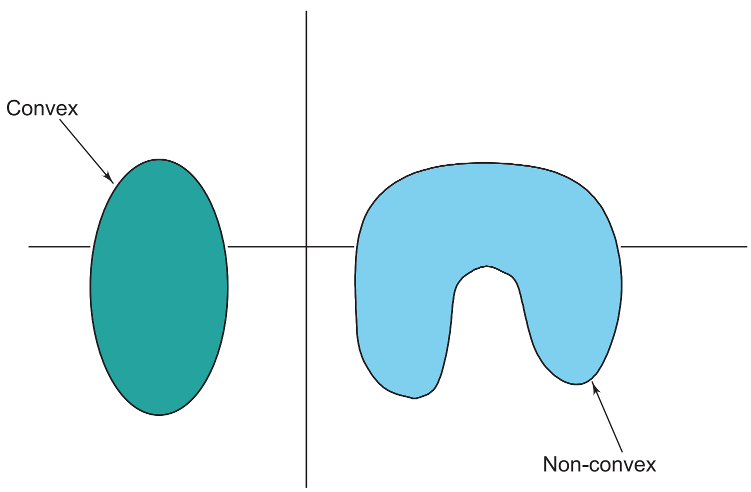

Fig. 4.2 Convex and non-convex sets in $\mathbb{R}^2$

Definition 4.2.2

The convex hull, $C(\mathcal{X}_i)$, of a pattern set $\mathcal{X}_i$ is the smallest convex set in $\mathbb{R}^n$ which contains the set $\mathcal{X}_i$. Equivalently, consider every convex set $\mathcal{S}_\alpha$, such that $\mathcal{X}_i \subset \mathcal{S}_\alpha \subset \mathbb{R}^n$, $\alpha \in \mathcal{I}$, where $\mathcal{I}$ is an index set. Then the convex hull of $\mathcal{X}_i$, $C(\mathcal{X}_i) = \bigcap_{\alpha\in\mathcal{I}} \mathcal{S}_\alpha$.

This only means that we need to take the intersection of all $\mathbb{R}^n$ subsets that contain the pattern set in question. This intersection yields the smallest convex set in $\mathbb{R}^n$ that contains the pattern set. Figure 4.2 shows a computer generated convex hull of the Iris data [12], using MATLAB. Notice how the convex hull of *iris sestosa* is disjoint from the convex hulls of *iris versicolor* and *iris virginica*. However, the convex hulls of *iris versicolor* and *iris virginica* are not separable.

The problem of convex hull computation is an interesting area of research. Neural networks have been applied to this problem (see [432]).

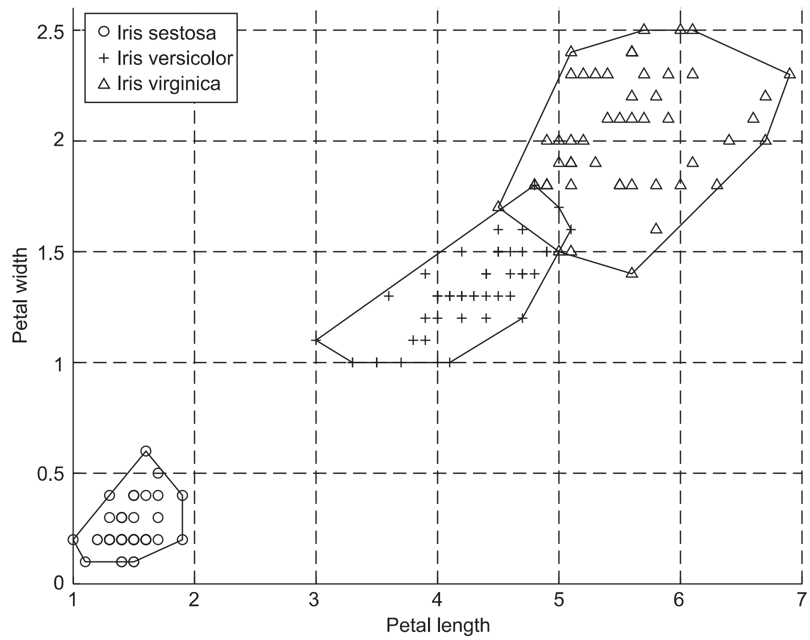

Fig. 4.3 MATLAB generated convex hulls for Iris data [12]

Note: This is a view of the 4-dimensional Iris data projected on to the x_{i3}–x_{i4} axes. The convex hull was generated using MATLAB's convhull command.

It should be intuitively clear that if the convex hulls of two pattern sets are non-overlapping, we can define a separating hyperplane that slices the pattern space into two halves, such that the pattern sets are separated. This leads to the concept of linear separability.

Definition 4.2.3

Two pattern sets $\mathcal{X}_i$ and $\mathcal{X}_j$ are said to be *linearly separable* if their convex hulls are disjoint, that is if $C(\mathcal{X}_i) \cap C(\mathcal{X}_j) = \phi$.

As shown in Fig. 4.3 since some minimum distance separates the two convex hulls, there exists a hyperplane Π, that separates the two sets. In the figures, we use the shorthand C_i to denote the convex hull $C(\mathcal{X}_i)$. The convex hull C_i is representative of the pattern class $\mathcal{C}_i$.

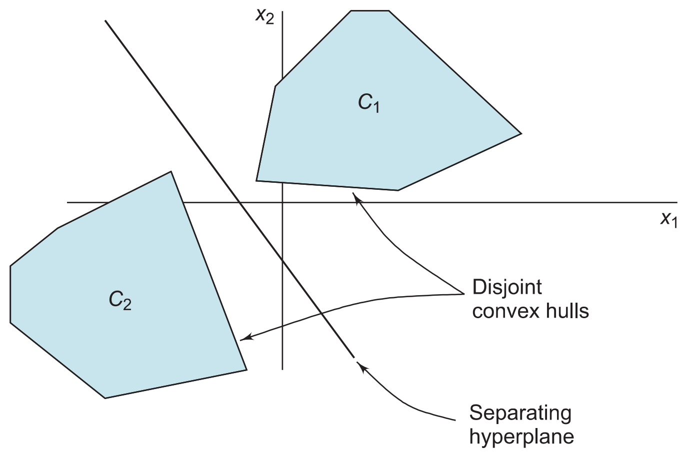

Fig. 4.4 Linearly separable pattern classes and a separating hyperplane

One separating hyperplane is the perpendicular bisector of the straight line joining the closest two points on the disjoint convex hulls.