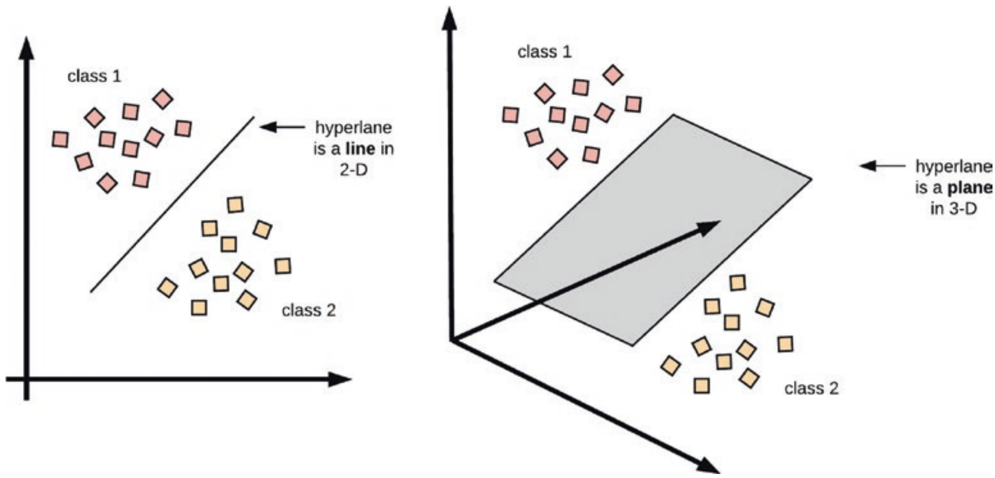

Figure 22-2. *Left: A hyperplane in 2-D is a line. Right: A hyperplane in 3-D is a plane. For dimension greater than 3, visualization becomes difficult.*

Finding the Optimal Hyperplane

The best hyperplane that linearly separates two classes is identified as the line lying at the largest margin from the nearest vectors at the boundary of the two classes.

In Figure 22-3, we observe that the best hyperplane is the line at the exact center of the two classes and constitutes the largest margin between both classes. Hence, this optimal hyperplane is also known as the largest margin classifier.

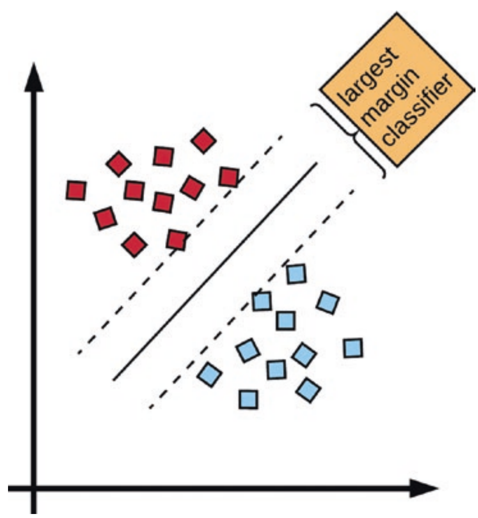

Figure 22-3. *The largest margin classifier*

The boundary points of the respective classes which are known as the support vectors are essential in finding the optimal hyperplane. The support vectors are illustrated in Figure 22-4. The boundary points are called support vectors because they are used to determine the maximum distance between the class they belong to and the discriminant function separating the classes.

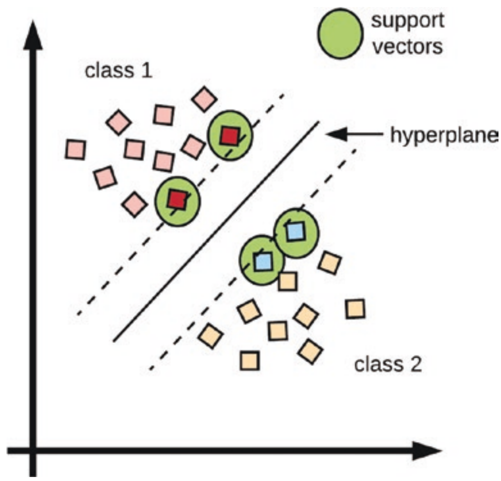

Figure 22-4. *Support vectors*

The mathematical formulation for finding the margin and consequently the hyperplane that maximizes the margin is beyond the scope of this book, but suffice to say this technique involves the Lagrange multiplier.

The Support Vector Classifier

In the real world, it is difficult to find data points that are precisely linearly separable and for which exists a large margin hyperplane. In Figure 22-5, the left image represents the data points for two classes in a dataset. Observe that there readily exists a linear separator between those two classes. Now, suppose we have an additional point from class 1 adjusted in such a way that it is much closer to class 2, we see that this point upsets the location of the hyperplane as seen in the right image of Figure 22-5. This reveals the sensitivity of the hyperplane to an additional data point that may result in a very narrow margin.