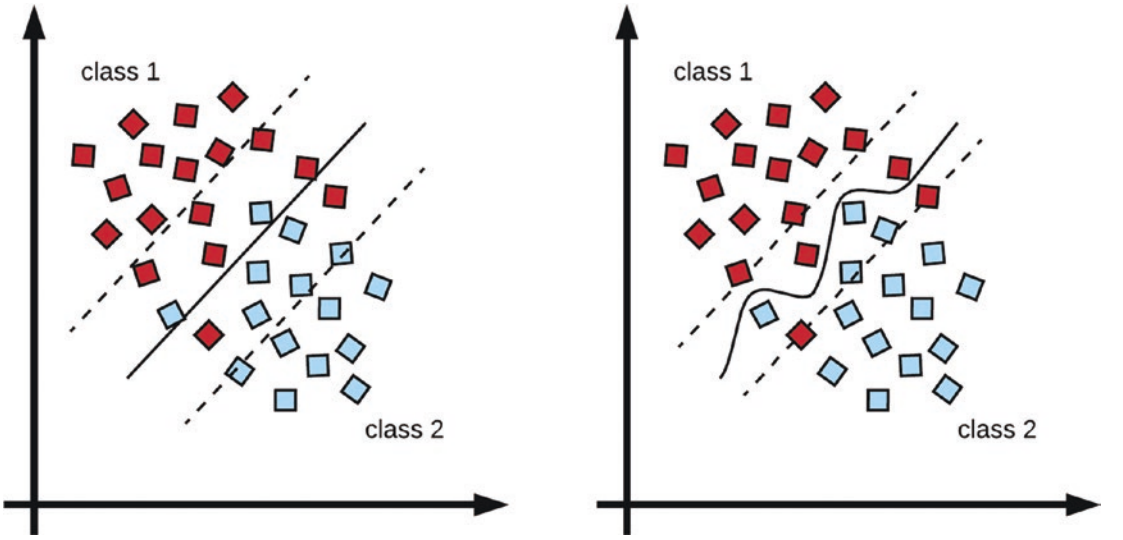

Figure 22-7. *Left: Higher values of C result in wider margins with more tolerance. Right: Lower values of C result in narrower margins with less tolerance*

Multi-class Classification

Previously, we have used the SVC to build a discriminant classifier for binary classes. What happens when we have more than two classes of outputs in the dataset, which is often the case in practice? The SVM can be extended for classifying k classes within a dataset, where k > 2. This extension is, however, not trivial with the SVM. There exist two standard approaches for addressing this problem. The first is the one-vs.-one (OVO) multi-class classification, while the other is the one-vs.-all (OVA) or one-vs.rest (OVR) multi-class classification technique.

One-vs.-One (OVO)

In the one-vs.-one approach, when the number of classes, k, is greater than 2, the algorithm constructs "k combination 2", $\left(\frac{k}{2}\right)$ classifiers, where each classifier is for a pair of classes. So if we have 10 classes in our dataset, a total of 45 classifiers is constructed or trained for every pair of classes. This is illustrated with four classes in Figure 22-8.

After training, the classifiers are evaluated by comparing examples from the test set against each of the $\left(\frac{k}{2}\right)$ classifiers. The predicted class is then determined by choosing the highest number of times an example is assigned to a particular class.