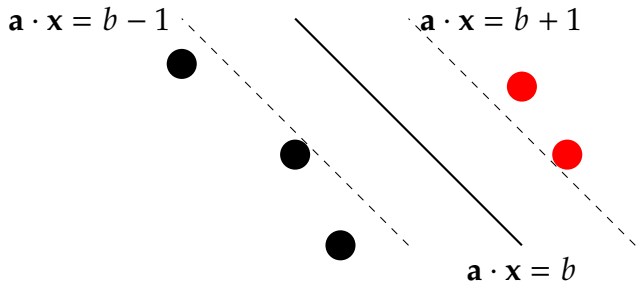
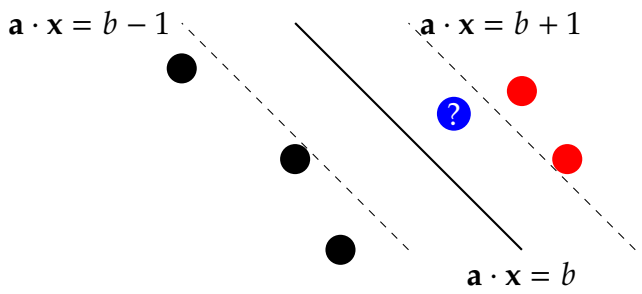
FAQs: Support Vector Machines

Why do we use $\mathbf{a} \cdot \mathbf{x} = b + 1$ when training the SVM hyperplane but $\mathbf{a} \cdot \mathbf{x} = b$ when classifying subsequent points?

Let's use the SVM to classify two types of points, indicated as red and black circles below. The SVM algorithm pushes the parallel $\mathbf{a} \cdot \mathbf{x} = b + 1$ and $\mathbf{a} \cdot \mathbf{x} = b - 1$ hyperplanes outward until they hit the closest training points (the support vectors).



Now that we've trained the SVM hyperplane, let's classify a new point. Imagine we want to classify the unknown (blue) point in the diagram below.



It seems clear that the blue point should be classified with the red ones. However, it is below the $\mathbf{a} \cdot \mathbf{x} = b + 1$ hyperplane but still above the $\mathbf{a} \cdot \mathbf{x} = b$ center hyperplane. We use the $\mathbf{a} \cdot \mathbf{x} = b$ hyperplane for classification to handle points that are inside the outermost hyperplanes.