CSE 250B: Section 5 - Sharad Vikram

1. SVM Decision Rule

   We have trained an SVM with an RBF kernel:

$$K(\mathbf{u}, \mathbf{v}) = \exp\left\{\frac{-(\mathbf{u}-\mathbf{v})^2}{2\sigma^2}\right\}$$

   Now we have a set of n support vectors (the training points the SVM keeps) $\{\mathbf{x}^{(i)}\}$, the associated training labels $\{y^{(i)}\}$ and alpha weights $\{\alpha_i\}$.

   What is the decision function $h(\mathbf{x})$.

Solution:

$$\begin{aligned} h(\mathbf{x}) &= \text{sign}(y_i \mathbf{w}^T \Phi(\mathbf{x})), \text{ and } \mathbf{w} = \sum_{i=1}^{n} \alpha_i \Phi(\mathbf{x}^{(i)}) \\ &= \text{sign}\Big(\sum_{i=1}^{n} \alpha_i y_i \Phi(\mathbf{x}^{(i)})^T \Phi(\mathbf{x})\Big) \\ &= \text{sign}\Big(\sum_{i=1}^{n} \alpha_i y_i K(\mathbf{x}^{(i)}, \mathbf{x})\Big) \\ &= \text{sign}\left(\sum_{i=1}^{n} \alpha_i y_i \exp\left\{\frac{-(\mathbf{x}^{(i)}-\mathbf{x})^2}{2\sigma^2}\right\}\right) \end{aligned}$$

2. Slack

   The "soft-margin" SVM includes slack for each variable, allowing it to fit non-linearly separable boundaries, as opposed to the "hard-margin" SVM, which does not have slack.

   Consider the soft-margin formulation:

$$\min \frac{1}{2}||w||^2 + C\sum_{i=1}^{N} \xi_i$$

   such that $\forall i$

$$y_i(w^T x + b) \geq 1 - \xi_i$$
$$\xi_i \geq 0$$

   What happens to the width of the margin in the following scenarios? Why?

   (a) $C \to 0$

   Solution: The width of the margin will go to infinity.

   (a) $C \to \infty$

Solution: The width of the margin will grow smaller until it misclassifies the least number of points it can.