

Homework

- Show that, irrespective of the dimensionality of the data space, a data set consisting of just two data points (call them $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$, one from each class) is sufficient to determine the maximum-margin hyperplane. Fully explain your answer, including giving an explicit formula for the solution to the hard margin SVM (i.e., \mathbf{w}) as a function of $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$.
- Gaussian kernel takes the form:

$$k(\mathbf{x}, \mathbf{x}') = \exp(-\|\mathbf{x} - \mathbf{x}'\|^2 / 2\sigma^2)$$

Try to show that the Gaussian kernel can be expressed as the inner product of an infinite-dimensional feature vector.

Hint: Making use of the following expansion, and then expanding the middle factor as a power series.

$$k(\mathbf{x}, \mathbf{x}') = \exp(-\mathbf{x}^T \mathbf{x} / 2\sigma^2) \exp(\mathbf{x}^T \mathbf{x}' / \sigma^2) \exp(-(\mathbf{x}')^T \mathbf{x}' / 2\sigma^2)$$