CS 536 : Support Vector Machine Problems

16:198:536

1) Suppose you had a data set in two dimensions that satisfied the following: the positive class all lay within a certain radius of a point, the negative class all lay outside that radius.

   - Show that under the feature map $\phi(x_1, x_2) = (1, x_1, x_2, x_1 x_2, x_1^2, x_2^2)$ (or equivalently, with the kernel $K(\underline{x}, \underline{y}) = (1 + \underline{x}.\underline{y})^2$), a linear separator can always be found in this embedded space, *regardless of radius and where the data is centered.*
   - In fact show that if there is an ellipsoidal separator, regardless of center, width, orientation (and dimension!), a separator can be found in the quadratic feature space using this kernel.

2) As an extension of the previous problem, suppose that the two dimensional data set satisfied the following: the positive class lay within one of two (disjoint) ellipsoidal regions, and the negative class was everywhere else. Argue that the kernel $K(\underline{x}, \underline{y}) = (1 + \underline{x}.\underline{y})^4$ will recover a separator.

3) Suppose that the two dimensional data set is distributed like the following: the positive class lays in a circle centered at some point, the negative class lies in a circular band surrounding it of some radius, and then additional positive points lie outside that radius. Argue that the kernel $K(\underline{x}, \underline{y}) = (1 + \underline{x}.\underline{y})^4$ will recover a separator.

4) Consider the XOR data (located at $(\pm 1, \pm 1)$). Express the dual SVM problem and show that a separator can be found using

   - $K(\underline{x}, \underline{y}) = (1 + \underline{x}.\underline{y})^2$
   - $K(\underline{x}, \underline{y}) = \exp(-||\underline{x} - \underline{y}||^2)$.

   For each, determine the regions of (x_1, x_2) space where points will be classified as positive or negative. Given that each produces a distinct separator, how might you decide which of the two was preferred?