SVM

1. a certain radius of a point.

A. K(x,y) = (1+x.y)2

Say the point is (x_{10}, x_{20}) and radius is r.
We have the boundary $(x_1 - x_{10})^2 + (x_2 - x_{20})^2 = r^2$.

Namely, the line separator is:

$$\begin{aligned} f(x_1, x_2) &= k((x_1 - x_{10})^2 + (x_2 - x_{20})^2 - r^2) \\ &= k(x_{10}^2 + x_{20}^2 - r^2)1 - 2kx_{10}x_1 - 2kx_{20}x_2 + kx_1^2 + kx_2^2 \end{aligned}$$

Therefore, for the given feature space $(1, x_1, x_2, x_1x_2, x_1^2, x_2^2)$,
$w = k(x_{10}^2 + x_{20}^2 - r^2, -2x_{10}, -2x_{20}, 0, 1, 1)$, where k is a parameter to control the margin, s.t. for all data points we have $|f(x_1, x_2)| \geq 1$.

B. an ellipsoidal separator

Notice that an ellipsoidal boundary can be represented as:

> $(x - b)^T C(x - b) = 1$, *where C is both symmetric and positive definite (for a non-degenerate d-dimensional ellipsoid in d dimensions).* [1]

Therefore, we have $y = k(\sum_{i=1}^{d} \sum_{j=1}^{d} c_{ij}(x_i - b_i)(x_j - b_j) - 1)$ to be the separator.

Obviously, all x terms shows in the form of $\{1, x_1, ..., x_d, x_1x_2, x1x_3, ..., x_{d-1}x_d, x_1^2, ..., x_d^2\}$. Namely, we can find a linear separator in this quadratic feature space using this kernel.

2. one of two (disjoint) ellipsoidal regions

Say we have a $f_1(x_1, x_2) = w_{10} + w_{11}x_1 + w_{12}x_2 + w_{13}x_1x_2 + w_{14}x_1^2 + w_{15}x_2^2$ as a separator for classifying one ellipsoidal region, and $f_2(x_1, x_2)$ for another region. We know all w's because of the proof of Q1.

We have $g(x_1, x_2)$ that can correctly separate the data set.

$$
\begin{aligned}
g(x_1, x_2) &= -\ f_1(x_1, x_2) f_2(x_1, x_2) \\
&= -\ w_{10} w_{20} - (w_{11} w_{20} + w_{10} w_{21}) x_1 - (w_{12} w_{20} + w_{10} w_{22}) x_2 \\
&\quad - (w_{13} w_{20} + w_{10} w_{23} + w_{11} w_{22} + w_{12} w_{21}) x_1 x_2 \\
&\quad - (w_{10} w_{24} + w_{14} w_{20} + w_{11} w_{21}) x_1^2 - (w_{10} w_{25} + w_{15} w_{20} + w_{12} w_{22}) x_2^2 \\
&\quad - (w_{11} w_{23} + w_{13} w_{21} + w_{14} w_{22} + w_{12} w_{24}) x_1^2 x_2 \\
&\quad - (w_{12} w_{23} + w_{13} w_{22} + w_{15} w_{21} + w_{11} w_{25}) x_1 x_2^2 \\
&\quad - (w_{11} w_{24} + w_{14} w_{21}) x_1^3 - (w_{12} w_{25} + w_{15} w_{22}) x_2^3 \\
&\quad - (w_{13} w_{24} + w_{14} w_{23}) x_1^3 x_2 \\
&\quad - (w_{14} w_{25} + w_{15} w_{24} + w_{13} w_{23}) x_1^2 x_2^2 \\
&\quad - (w_{13} w_{25} + w_{15} w_{23}) x_1 x_2^3 \\
&\quad - w_{14} w_{24} x_1^4 - w_{15} w_{25} x_2^4
\end{aligned}
$$

Say the equivalent feature map

$\phi(x_1, x_2) = (1, x_1, x_2, x_1x_2, x_1^2, x_2^2, x_1^2x_2, x_1x_2^2, x_1^3, x_2^3, x_1x_2^3, x_1^2x_2^2, x_1x_2^3, x_1^4, x_2^4)$.

(Actually, some coefficients are not 1 if we use the kernel directly. But for simplicity, assume we normalize it so that all coefficients are 1.)

We have:

$$
\begin{aligned}
w = -1(&w_{10} w_{20}, w_{11} w_{20} + w_{10} w_{21}, \\
&w_{13} w_{20} + w_{10} w_{23} + w_{11} w_{22} + w_{12} w_{21}, \\
&w_{10} w_{24} + w_{14} w_{20} + w_{11} w_{21}, w_{10} w_{25} + w_{15} w_{20} + w_{12} w_{22}, \\
&w_{11} w_{23} + w_{13} w_{21} + w_{14} w_{22} + w_{12} w_{24}, w_{12} w_{23} + w_{13} w_{22} + w_{15} w_{21} + w_{11} w_{25}, \\
&w_{11} w_{24} + w_{14} w_{21}, w_{12} w_{25} + w_{15} w_{22}, \\
&w_{13} w_{24} + w_{14} w_{23}, w_{14} w_{25} + w_{15} w_{24} + w_{13} w_{23}, w_{13} w_{25} + w_{15} w_{23}, \\
&w_{14} w_{24}, w_{15} w_{25})
\end{aligned}
$$

Therefore, we can find the linear separator with the kernel $K(x, y) = (1 + x.y)^4$

3. circle, circular band, outside

Say the center is (x_{10}, x_{20}), inner radius of the band is r_1, and outer radius is r_2.

We have a $f_1(x_1, x_2) = w_{10} + w_{11}x_1 + w_{12}x_2 + w_{13}x_1x_2 + w_{14}x_1^2 + w_{15}x_2^2$ that can separate the inner circle and the band, and $f_2(x_1, x_2)$ that can separate the band and outside. We know all w's because of the proof of Q1.

(Say f_1, f_2 will return positive values for inner points and neagtive values for outer ones.)

We have $g(x_1, x_2) = f_1(x_1, x_2) f_2(x_1, x_2)$

For the same reason as Q2,

$$\begin{aligned} w = (&w_{10}w_{20},\ w_{11}w_{20} + w_{10}w_{21}, \\ &w_{13}w_{20} + w_{10}w_{23} + w_{11}w_{22} + w_{12}w_{21}, \\ &w_{10}w_{24} + w_{14}w_{20} + w_{11}w_{21},\ w_{10}w_{25} + w_{15}w_{20} + w_{12}w_{22}, \\ &w_{11}w_{23} + w_{13}w_{21} + w_{14}w_{22} + w_{12}w_{24},\ w_{12}w_{23} + w_{13}w_{22} + w_{15}w_{21} + w_{11}w_{25}, \\ &w_{11}w_{24} + w_{14}w_{21},\ w_{12}w_{25} + w_{15}w_{22}, \\ &w_{13}w_{24} + w_{14}w_{23},\ w_{14}w_{25} + w_{15}w_{24} + w_{13}w_{23},\ w_{13}w_{25} + w_{15}w_{23}, \\ &w_{14}w_{24},\ w_{15}w_{25}) \end{aligned}$$

Therefore, we can find the linear separator with the kernel $K(x, y) = (1 + x.y)^4$

4. XOR data

A. dual problem

Find the set of values $\{a_i\}$ that solves the following:

$$\max_a \sum_{i=1}^{m} a_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{i=1}^{m} a_i y^i K(x^i, x^j) y^j a_j,\ s.t. \sum_{i=1}^{m} a_i y^i = 0, \forall i, a_i \geq 0.$$

Here we have $m = 4$:

$$\max_a \sum_{i=1}^{4} a_i - \frac{1}{2}\sum_{i=1}^{4}\sum_{i=1}^{4}(-1)^{i+j} a_i K(x^i, x^j) a_j,\ s.t. a_1 + a_3 = a_2 + a_4, a_i \geq 0$$

B. K(x,y) = (1 + x.y)2

Here we have $\max_a \sum_{i=1}^{4} a_i - \frac{1}{2}\sum_{i=1}^{4}\sum_{i=1}^{4}(-1)^{i+j} a_i a_j (1 + x^i.x^j)^2$

Let $I = \sum_{i=1}^{4} a_i - \frac{1}{2}\sum_{i=1}^{4}\sum_{i=1}^{4}(-1)^{i+j} a_i a_j (1 + x^i.x^j)^2$

We have $I = \sum_{i=1}^{4}(a_i - \frac{9}{2}a_i^2) + a_1a_2 - a_1a_3 + a_1a_4 + a_2a_3 - a_2a_4 + a_3a_4.$

Here we have $a_i = \frac{1}{8}, i = 1, 2, 3, 4$ maximize I given the constrains.

Therefore, we have $y = \sum_{i=1}^{4} a_i y^i (1 + x^i x)^2 = -x_1x_2$ as the separator.

C. K(x,y) = exp(−||x−y||2).

Here we have $\max_a \sum_{i=1}^{4} a_i - \frac{1}{2}\sum_{i=1}^{4}\sum_{i=1}^{4}(-1)^{i+j}a_i a_j \exp(-||x^i - x^j||^2)$

Let $I = \sum_{i=1}^{4} a_i - \frac{1}{2}\sum_{i=1}^{4}\sum_{i=1}^{4}(-1)^{i+j}a_i a_j \exp(-||x^i - x^j||^2)$

We have

$$I = \sum_{i=1}^{4}(a_i - \frac{1}{2}a_i^2) + (a_1a_2 + a_1a_4 + a_2a_3 + a_3a_4)\exp(-4) - (a_1a_3 + a_2a_4)\exp(-8)$$

Here we have $a_i = \frac{1}{(1-\exp(-4))^2}$, $i = 1, 2, 3, 4.$

Therefore, $y = \sum_{i=1}^{4} a_i y^i \exp(-||x^i - x||^2) = \frac{1}{1-\exp(-4))^2}\sum_{i=1}^{4}(-1)^i \exp(-||x^i - x||^2).$

D. Preference

Notice that those 2 separators have the exactly same boundary: $x_1 = 0, x_2 = 0$(, and therefore the same results. More specifically, the first and third quadrants are negative and the second and forth are possitive). But the Gussian Kernel is much harder to compute and sort of hard to interpret (to humans). The polynomial kernel is much simpler in terms of computing and the output separator, and the separator is so simple that we can directly get its idea: “if x_1 and x_2 have the same sign, $y = -1$, and vice versa.”, which is just the meaning of XOR.

(Nevertheless, this conclusion is conditioned on the XOR data. If the data is much more harder to separate, which might be so hard that we cannot tell the oreder of the polynomial kernel intuitively, we might have to try a large number of kernels. In this case, Gussian Kernel might be helpful.)

Another point is, the Gussian Kernel is more complex than polynomial ones, and therefore it tends to overfitting when there is noise, which is common in practice.

---

1. MATHEMATICAL REPRESENTATION OF AN ELLIPSOID: https://www.physics.smu.edu/~scalise/SMUpreprints/SMU-HEP-10-14.pdf ↩