Problem 1

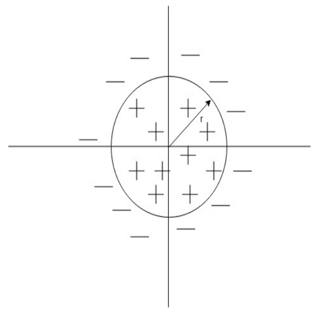

Figure 1: Problem 1.1

The problem states that we have a dataset in two dimensions that looks something like 1.We have all the positive class values inside a circle of radius "r" and all the negative class values are outside of the circle. Looking at the data, it is clear that this data does not have a linear separator in 2 dimensions of features. So, we need to come up with a higher dimensional feature space which can possibly have a linear separator. Looking at the image, we can decipher few things:

$$T = r^2 - X_1^2 - X_2^2$$
$$classify(\underline{X}) = sign(T)$$
$$= \begin{cases} +1, & \text{if } T \geq 0 \\ -1, & T < 0 \end{cases}$$

From the above equations, we see that if somehow we have data transformed to a higher dimension which has the squares of the features, we can successfully find a linear separator. This is possible by having a mapping function $\phi(x_1, x_2) = (1, x_1, x_2, x_1x_2, x_1^2, x_2^2)$. T can be obtained from this mapping function by keeping coefficients of parameters as $1 : r^2, x_1^2 : -1, x_2^2 : -1$.

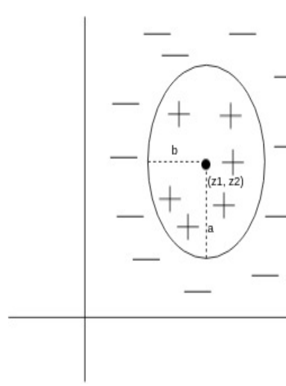

Figure 2: Problem 1.2

For the second part 2, we need to find an ellipsoidal separator, regardless of center, width etc. Let us try to expand an equation for the ellipse:

$$
\begin{aligned}
T &= b^2(x_1 - z_1)^2 + a^2(x_2 - z_2)^2 - a^2b^2 \\
classify(\underline{X}) &= sign(T) \\
&= \begin{cases} +1, & \text{if } T \leq 0 \\ -1, & T > 0 \end{cases}
\end{aligned}
$$

z_1, z_2 are centers of ellipse. We see that the parameters of this equation can be achieved by the same mapping function.

Problem 2

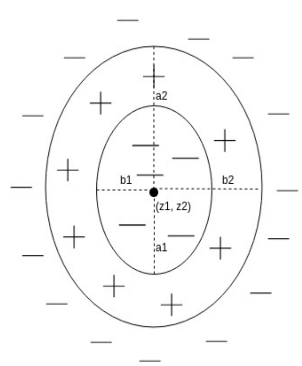

Figure 3: Problem 2

This problem can be represented by the image above 3. In order to find a linear separator for this case, we will have to find one in higher dimensional feature space. Let us assume inner ellipse's parameters to be a_1 and b_1, outer one's to be a_2 and b_2. Also, label the three regions as 1, 2, and 3 (1 and 3 are negative class, 2 is positive class). Using the ellipsoidal equations from previous question, we can represent two functions as:

$$
\begin{aligned}
T_1 &= b_1^2(x_1 - z_1)^2 + a_1^2(x_2 - z_2)^2 - a_1^2b_1^2 \\
T_2 &= b_2^2(x_1 - z_1)^2 + a_2^2(x_2 - z_2)^2 - a_2^2b_2^2
\end{aligned}
$$

For 1:

$$
\begin{aligned}
T_1 &< 0 \\
T_2 &< 0
\end{aligned}
$$

For 2:

$$
\begin{aligned}
T_1 &> 0 \\
T_2 &< 0
\end{aligned}
$$

For 3:

$$T_1 > 0$$
$$T_2 > 0$$

Now, we can make a separator by multiplying $T_1 * T_2$. By doing this, we get:

$$\begin{aligned} classify(\underline{X}) &= sign(T_1 * T_2) \\ &= \begin{cases} -1, & \text{if } (T_1 * T_2) \geq 0 \\ +1, & (T_1 * T_2) < 0 \end{cases} \end{aligned}$$

$$\begin{aligned} T_1 * T_2 &= (b_1^2(x_1 - z_1)^2 + a_1^2(x_2 - z_2)^2 - a_1^2 b_1^2)(b_2^2(x_1 - z_1)^2 + a_2^2(x_2 - z_2)^2 - a_2^2 b_2^2) \\ &= (b_1^2 x_1^2 + b_1^2 z_1^2 - 2b_1^2 x_1 z_1 + a_1^2 x_2^2 + a_1^2 z_2^2 - 2a_1^2 x_2 z_2 - a_1^2 b_1^2) \\ &(b_2^2 x_1^2 + b_2^2 z_1^2 - 2b_2^2 x_1 z_1 + a_2^2 x_2^2 + a_2^2 z_2^2 - 2a_2^2 x_2 z_2 - a_2^2 b_2^2) \end{aligned}$$

On expanding $T_1 * T_2$, we get the parameters $1, x_1, x_2, x_1x_2, x_1^2, x_2^2, x_1^2x_2^2 .... x_1^4, x_2^4$, which can be represented by the mapping function:

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

. We can now use the kernel function to represent this high dimensional mapping function through:

$$K(\underline{x}, \underline{y}) = (1 + \underline{x}.\underline{y})^4$$

.

Problem 3

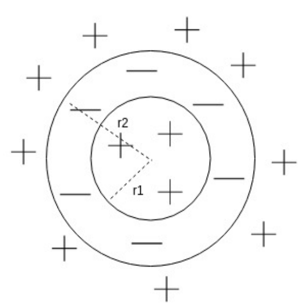

Figure 4: Problem 3

Looking at the image 4, we see that the negative class points are sandwiched between positive class points. In order to find a linear separator for this case, we will have to find one in higher dimensional feature space. Let us assume inner circle's radius to be r_1 and outer one's to be r_2.

Also, label the three regions as 1, 2, and 3 (1 and 3 are positive class, 2 is negative class). Let there be two equations:

$$T_1 = r_1^2 - X_1^2 - X_2^2$$
$$T_2 = r_2^2 - X_1^2 - X_2^2$$

For 1:

$$T_1 > 0$$
$$T_2 > 0$$

For 2:

$$T_1 < 0$$
$$T_2 > 0$$

For 3:

$$T_1 < 0$$
$$T_2 < 0$$

Now, we can make a separator by multiplying $T_1 * T_2$. By doing this, we get:

$$\begin{aligned} classify(\underline{X}) &= sign(T_1 * T_2) \\ &= \begin{cases} +1, & \text{if } (T_1 * T_2) \geq 0 \\ -1, & (T_1 * T_2) < 0 \end{cases} \end{aligned}$$

$$\begin{aligned} T_1 * T_2 &= (r_1^2 - X_1^2 - X_2^2)(r_2^2 - X_1^2 - X_2^2) \\ &= x_1^4 + x_2^4 + 2x_1^2x_2^2 - x_1^2(r_1^2 + r_2^2) - x_2^2(r_1^2 + r_2^2) + r_1^2r_2^2 \end{aligned}$$

On expanding $T_1 * T_2$, we will see that this higher dimensional feature space can be expressed by the kernel $K(\underline{x}, \underline{y}) = (1 + \underline{x}.\underline{y})^2$.

Problem 4

The data can be represented by the below table:

x	Co-ordinates	Result
x_1	(-1, -1)	-1
x_2	(-1, +1)	+1
x_3	(+1, -1)	+1
x_4	(+1, +1)	-1

In the first part, the kernel function is given as $K(\underline{x}, \underline{y}) = (1 + \underline{x}.\underline{y})^2$. We can calculate the values of this kernel matrix in the following way:

$$\begin{aligned} K_{11} &= (1 + \begin{bmatrix} -1 \\ -1 \end{bmatrix} \begin{bmatrix} -1 & -1 \end{bmatrix}) \\ &= (1 + 2)^2 \\ &= 9 \end{aligned}$$

Doing calculations like this, we get the final K-matrix as: $\begin{bmatrix} 9 & 1 & 1 & 1 \\ 1 & 9 & 1 & 1 \\ 1 & 1 & 9 & 1 \\ 1 & 1 & 1 & 9 \end{bmatrix}$ Multiplying this matrix with the result vector, we get K' = $\begin{bmatrix} 9 & -1 & -1 & 1 \\ -1 & 9 & 1 & -1 \\ -1 & 1 & 9 & -1 \\ 1 & -1 & -1 & 9 \end{bmatrix}$

Now, according to Dual SVM, and taking its derivative

$$L(\alpha) = \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i y^i K_{ij} y^j \alpha_j$$
$$\nabla(L) = 1 - K\alpha$$
$$\alpha = \begin{bmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$$

These values for α satisfy the conditions imposed on dual SVM equation, so we can say that it is a valid solution. Now, we need to find $w = \sum_i \alpha_i y_i \phi(x_i)$, using the feature mapping function as $\phi(x_1, x_2) = (1, x_1, x_2, x_1x_2, x_1^2, x_2^2)$ which comes out to be:

$$w = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, we have to find the separating function as $y = w^T\phi(x) = -x_1x_2$.

For the second part, the kernel function is given as $K(\underline{x}, \underline{y}) = \exp\left(-(||\underline{x} - \underline{y}||)^2\right)$. We can calculate the values of this kernel matrix in the following way:

$$||\underline{x} - \underline{y}|| = (x_1 - y_1)^2 + (x_2 + y_2)^2$$
$$K_{11} = e^{-((-1+1)^2+(-1+1)^2)}$$
$$= 1$$
$$K_{12} = e^{-((-1+1)^2+(-1-11)^2)}$$
$$= e^{-4}$$

Doing calculations like this, we get the final K-matrix as:

$$\begin{bmatrix} 1 & e^{-4} & e^{-4} & e^{-8} \\ e^{-4} & 1 & e^{-8} & e^{-4} \\ e^{-4} & e^{-8} & 1 & e^{-4} \\ e^{-4} & e^{-4} & e^{-4} & 1 \end{bmatrix}$$

Multiplying this matrix with the result vector, we get K':

$$\begin{bmatrix} 1 & -e^{-4} & -e^{-4} & e^{-8} \\ -e^{-4} & 1 & e^{-8} & -e^{-4} \\ -e^{-4} & e^{-8} & 1 & -e^{-4} \\ e^{-4} & -e^{-4} & -e^{-4} & 1 \end{bmatrix}$$

Now, according to Dual SVM, and taking its derivative

$$\begin{aligned} L(\alpha) &= \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i y^i K_{ij} y^j \alpha_j \\ \nabla(L) &= 1 - K\alpha \\ \alpha &= \begin{bmatrix} e^{0.03} & e^{0.03} & e^{0.03} & e^{0.03} \end{bmatrix} \end{aligned}$$

These values for α satisfy the conditions imposed on dual SVM equation, so we can say that it is a valid solution. Now, we need to find $w = \sum_i \alpha_i y_i \phi(x_i)$, but since feature mapping function for Gaussian kernel is infinite, I am not able to calculate further. I will try to complete and submit it again, if allowed. I believe with some more time, I should be able to figure it out.