**Example:2 (Illustration of training a nonlinear SVM)**

Consider the XOR problem:

|  |  |
| --- | --- |
| Input Vector | Desired output y |
|  |  |
|  |  |
|  |  |
|  |  |



Figure 13: square showing the negative and circles showing the positive examples

As shown in the above figure, there is no linear function that can separate the four training points. So, we will use the kernel function as

(a)

Where, =( and =(

The kernel function can be expressed as

) (b)

The images of input vector is induced in the feature space is therefore deduced to be

) (c )

The dual form of the maximization function is

W()= - (d)

Here, all four points are the Support Vectors and by replacing the kernel function in equation (a), we will have

W()= – (e)

= ++ (f)

= 2

Similarly,

=0 = -2 =2 = -2

=0 =2 =0 =2

Substituting these values in equation (f), we will get

W() = ++ (g) Optimizing W() with respect to the Lagrange multipliers yields the following set of simultaneous equations:

w.r.to : 1 -

- (h)

w.r.to : + (i)

w.r.to : +9 (j)

w.r.to : - (k)

Solving equation (h) and (k),we get

= (l)

Solving equation (i) and (j), we get

= (m)

By putting the value for and in equation (h),

– (n)

Solving equation (h) and (i), we have

+ (o)

Multiply equation (n) with 2 and add with (0), we get

= = (p)

Putting value of in equation (o), and solving for

= = (q)

By putting this values of in equation (g), the optimum value of W() is

W()= i.e 2 = or =

Now, to find the optimal hyperplane we need to find

=

= +

=

= =

The bias b is 0, because the first element of is 0.

The optimal hyperplane becomes

) + b = 0

Substituting the values of and b,

[0 0 0 0] = 0

Which reduces to optimal hyperplane,