Mid2 – 2018 paper

tutorial

- **1.** Consider a two dimensional (input-space with co-ordinates $(x_1, x_2)^t$) two class problem. Training set for $+1\ class$ is $\{X_1 = (0,1)^t, X_2 = (0,-1)^t, X_3 = (1,0)^t, X_4 = (-1,0)^t\}$; that for -1 class is $\{X_5 = (0,0)^t\}$. It is known that in the feature-space with co-ordinates $\left(x_1^2, x_2^2,\ \sqrt{2}\, x_1 x_2\right)$ the given problem is a linearly separable one. Can you apply *perceptron learning with single sample correction* (consider the training patterns in the order $X_1, X_2, \ldots, X_5$) in the feature-space and thus find a non-linear (quadratic) classifier in the input space. Let your initial solution be the Zero Vector (in the feature-space). Your final answer should be a classifier in the input space (quadratic equation in the input space) which you can show geometrically (pictorially) also {Show necessary intermediate steps}. [8 marks]

- This problem you can do.

- **2.** For the question 1, with kernel $k(X_i, X_j) = (X_i \cdot X_j)^2$ find the hard nonlinear SVM classifier. From the geometry of the problem, one can assume that Lagrange multipliers for all positive training examples are same, i.e., $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$. Draw the classifier in the input-space. Can you find the margin in the feature-space? [9 marks]

Class +1:

x1 = (0,1) , x2 = (0,-1), x3 = (1,0),x4 = (-1,0)

Class -1:

X5 = (0,0).

Class +1:

x1 = (0,1) , x2 = (0,-1), x3 = (1,0),x4 = (-1,0)

Class -1:

X5 = (0,0).

Kernel function : $k(X_i, X_j) = (X_i \cdot X_j)^2$.

Given lagrange multipliers :α1 = α2= α3= α4 = α. The Lagrange for fifth co-ordinate be α5.

Wolf dual :

L(α) = $\sum_{i=1}^{5} \alpha_i - \frac{1}{2} \sum_i^5 \sum_j^5 \alpha_i \alpha_j y_i y_j . (X_{i.} X_j)$

(few students did without considering α5, this time we considered it)

L = $\sum_{i=1}^{5} \alpha_i - \frac{1}{2} \sum_{i}^{5} \sum_{j}^{5} \alpha_i \alpha_j y_i y_j . K(X_{i.} X_j)$ (in feature space)

Condition: $\sum_{i=1}^{5} \alpha_i . y_i$ = 0 && α_i >= 0, i = 1,2,3,4,5

$$L = \sum_{i=1}^{5} \alpha_i - \frac{1}{2}\sum_{i}^{5}\sum_{j}^{5} \alpha_i\alpha_j y_i y_j . K(X_{i.}X_j) \text{ (in feature space)}$$

Condition: $\sum_{i=1}^{5} \alpha_i \, . \, y_i = 0$ && $\alpha_i >= 0$, i = 1,2,3,4,5

Table 1:

$K(X_{i.}X_j)$	X_1	X_2	X_3	X_4	X_5
X_1	1	1	0	0	0
X_2	1	1	0	0	0
X_3	0	0	1	1	0
X_4	0	0	1	1	0
X_5	0	0	0	0	0

Table 2:

	y_1	y_2	y_3	y_4	y_5
y_1	1	1	1	1	-1
y_2	1	1	1	1	-1
y_3	1	1	1	1	-1
y_4	1	1	1	1	-1
y_5	-1	-1	-1	-1	1

$= \sum_{i=1}^{5} \alpha_i . y_i = 0 =>$ $(1^* \alpha + 1^* \alpha + 1^* \alpha + 1^* \alpha - 1^* \alpha_5) = 4\alpha - \alpha_5 = 0$. Therefore $\alpha_5 = 4\alpha$.

L(α) = $\sum_{i=1}^{5} \alpha_i - \frac{1}{2}\sum_{i}^{5}\sum_{j}^{5} \alpha_i\alpha_j y_i y_j . K(X_{i.}X_j)$ (in feature space)

$4\alpha + 4\alpha\ - \frac{1}{2}(\alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1)$. Based on table 2. Ignoring the combination which results in 0.

L(α) = $\sum_{i=1}^{5} \alpha_i - \frac{1}{2} \sum_i^5 \sum_j^5 \alpha_i \alpha_j y_i y_j . K(X_{i.} X_j)$ (in feature space)

$4\alpha + 4\alpha - \frac{1}{2}(\alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1 + \alpha^2 * 1 * 1)$. Based on table 2. Ignoring the combination which results in 0.

L(α) = $8\alpha - \frac{1}{2}(8 * \alpha^2)$

$\frac{\partial L}{\partial \alpha} = 0 => 8 - 8\alpha = 0 => \alpha = 1$

From question: $\alpha_{1=} \alpha_{2=} \alpha_{3=} \alpha_{4=} \alpha$ = 1 . $\alpha_{5=} 4\alpha = 4$

b = y_j - $\sum_{i=1}^{5} \alpha_i . y_i . (X_i . X_j)$. j can be selected any value from 1 to 5. Let j = 1

b = 1 – (1*1*1+1*1*1) (* Based on table 2, for 1 only 1 and 2 give non zero values).

B = 1-2 = -1.

Classifier :

g(x) = $\sum_{i=1}^{5} \alpha_i . y_i . K(X_i . X_j)$+b = 0

= α * $\sum_{i=1}^{5} y_i . K(X_i . X)$-1=0

Computing $K(X_i . X)$: I = 1,2,3,4,5. x1 = (0,1) , x2 = (0,-1), x3 = (1,0),x4 = (-1,0),x5= (0,0)

K($X_1 . X$) = ([0,1] [x1 x2]^T)^2 = x2^2 .

K($X_2 . X$) = ([0,-1] [x1 x2]^T)^2 = (-x2)^2 = x2^2

K($X_3 . X$) = ([1,0] [x1 x2]^T)^2 = x1^2

K($X_4 . X$) = ([1,0] [x1 x2]^T)^2 = x1^2

K($X_5 . X$) = ([0,0] [x1 x2]^T)^2 = 0

g(x) = 1 * (1* x2^2+1* x2^2+1* x1^2+1* x1^2+0)-1 = 0

$$g(x) = 1 * (1* x2^2+1* x2^2+1* x1^2+1* x1^2+0)-1 = 0$$

$$= x2^2+ x2^2+ x1^2+ x1^2-1 = 0$$

$$= 2\, x2^2 + 2\, x1^2=1$$

$$x2^2 + x1^2 = ½.$$

Feature space co-ordinates = $\left(x_1^2, x_2^2,\ \sqrt{2}\, x_1 x_2 \right)$.

Normal Space X = (x1,x2)	Feature Space X = $\left(x_1^2, x_2^2,\ \sqrt{2}\, x_1 x_2 \right)$
X1 = $(0,1)$	X1 = $(0,1,0)$
X2 = $(0,-1)$	X2 = $(0,1,0)$
X3 = $(1,0)$	X3 = $(1,0,0)$
X4 = $(-1,0)$	X4 = $(1,0,0)$
X5 = (0,0)	X5 = (0,0,0)

W in feature space:

W = $\sum_{i=1}^{5} \alpha_i . y_i . \Theta(X_i . X_j)$ = 1 * 1 * $(0,1,0)^T$ + 1 * 1 * $(0,1,0)^T$+1 * 1 * $(1,0,0)^T$+ 1 * 1 * $(1,0,0)^T$+1 * 1 * $(0,0,0)^{T}$.

Feature space co-ordinates = $(x_1^2, x_2^2, \sqrt{2}\, x_1 x_2)$.

Normal Space X = (x1,x2)	Feature Space X = $(x_1^2, x_2^2, \sqrt{2}\, x_1 x_2)$
X1 = $(0,1)$	X1 = $(0,1,0)$
X2 = $(0,-1)$	X2 = $(0,1,0)$
X3 = $(1,0)$	X3 = $(1,0,0)$
X4 = $(-1,0)$	X4 = $(1,0,0)$
X5 = (0,0)	X5 = (0,0,0)

W in feature space:

W = $\sum_{i=1}^{5} \alpha_i . y_i . \Theta(X_i . X_j)$ = 1 * 1 * $(0,1,0)^T$ + 1 * 1 * $(0,1,0)^T$+1 * 1 * $(1,0,0)^T$+ 1 * 1 * $(1,0,0)^T$+1 * 1 * $(0,0,0)^T$.

= $(0,1,0)^T + (0,1,0)^T + (1,0,0)^T + (1,0,0)^T = (2\ 2\ 0)^T$

Margin = 2 / ||W|| = 2 / √(4+4) = 2 / 2√2 = 1/ √2