

# Introduction to Machine Learning

## Unit 8 Problems: Support Vector Machines

Prof. Sundeep Rangan

1. Consider the data set for four points with features  $\mathbf{x}_i = (x_{i1}, x_{i2})$  and binary class labels  $y_i = \pm 1$ .

$x_{i1}$	0	1	1	2
$x_{i2}$	0	0.3	0.7	1
$y_i$	-1	-1	1	1

- (a) Find a linear classifier that separates the two classes. Your classifier should be of the form

$$\hat{y} = \begin{cases} 1 & \text{if } b + w_1x_1 + w_2x_2 > 0 \\ -1 & \text{if } b + w_1x_1 + w_2x_2 < 0 \end{cases}$$

State the intercept  $b$  and weights  $w_1$  and  $w_2$  for your classifier. Note there is no unique answer as there are multiple linear classifiers that could separate the classes.

- (b) Find the maximum  $\gamma$  such that

$$y_i(b + w_1x_{i1} + w_2x_{i2}) \geq \gamma, \text{ for all } i,$$

for the classifier in part (a)?

- (c) Compute the margin of the classifier

$$m = \frac{\gamma}{\|\mathbf{w}\|}, \quad \|\mathbf{w}\| = \sqrt{w_1^2 + w_2^2}.$$

- (d) Which samples  $i$  are on the margin for your classifier?

2. Consider the data set with scalar features  $x_i$  and binary class labels  $y_i = \pm 1$ .

$x_i$	0	1.3	2.1	2.8	4.2	5.7
$y_i$	-1	-1	-1	1	-1	1

Consider a linear classifier for this data of the form,

$$\hat{y} = \begin{cases} 1 & z > 0 \\ -1 & z < 0, \end{cases} \quad z = x - t,$$

where  $t$  is a threshold. For each threshold  $t$ , let  $J(t)$  denote the sum hinge loss,

$$J(t) = \sum_i \epsilon_i, \quad \epsilon_i = \max(0, 1 - y_i z_i).$$

- (a) Write a short python program to plot  $J(t)$  vs.  $t$  for 100 values of  $t$  in the interval  $t \in [0, 5]$ .
- (b) Based on the plot, what is one value of  $t$  that minimizes  $J(t)$ .
- (c) For the value of  $t$  in part (b), find the corresponding slack variables  $\epsilon_i$ .
- (d) Which samples  $i$  violate the margin ( $\epsilon_i > 0$ ) and which samples  $i$  are misclassified ( $\epsilon_i > 1$ ).
3. Consider an image recognition problem, where an image  $\mathbf{X}$  and filter  $\mathbf{W}$  are  $4 \times 4$  matrices:

$$\mathbf{X} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- (a) Recall that in linear classification, the  $4 \times 4$  image matrices  $\mathbf{X}$  and  $\mathbf{W}$  can be represented as 16-dimensional vectors,  $\mathbf{x} = \text{vec}(\mathbf{X})$  and  $\mathbf{w} = \text{vec}(\mathbf{W})$  by stacking the columns of the matrices vertically. What are  $\mathbf{x}$  and  $\mathbf{w}$  for the matrices above.
- (b) What is the inner product  $z = \mathbf{w}^T \mathbf{x}$ .
- (c) What is the inner product  $z = \mathbf{w}^T \mathbf{x}_{\text{right}}$  where  $\mathbf{x}_{\text{right}}$  is the vector corresponding to the matrix  $\mathbf{X}$  right shifted by one pixel with the left column filled with zeros.
- (d) What is the inner product  $z = \mathbf{w}^T \mathbf{x}_{\text{left}}$  where  $\mathbf{x}_{\text{left}}$  is the vector corresponding to the matrix  $\mathbf{X}$  left shifted by one pixel with the right column filled with zeros.
- (e) Write the python command that can convert a  $4 \times 4$  image matrix, `Xmat` to the 16-dimensional vector, `x`. What is the python command to go from `x` to `Xmat`.
4. Consider the data set with scalar features  $x_i$  and binary class labels  $y_i = \pm 1$ .

$x_i$	0	1	2	3
$y_i$	1	-1	1	-1

A support vector classifier is of the form

$$\hat{y} = \begin{cases} 1 & z > 0 \\ -1 & z < 0, \end{cases} \quad z = \sum_i \alpha_i y_i K(x_i, x),$$

where  $K(x, x')$  is the radial basis function,  $K(x, x') = e^{-\gamma(x-x')^2}$ , and  $\gamma > 0$  and  $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_4]$  are parameters of the classifier.

- (a) Use python to plot  $z$  vs.  $x$  and  $\hat{y}$  vs.  $x$  when  $\gamma = 3$  and  $\boldsymbol{\alpha} = [0, 0, 1, 1]$ .
- (b) Repeat (a) with  $\gamma = 0.3$  and  $\boldsymbol{\alpha} = [1, 1, 1, 1]$ .
- (c) Which classifier makes more errors on the training data.