

1.

KKT for SVM

$$\mathbf{w} = \sum_{k=1}^K \lambda_k y_k \mathbf{x}_k$$

Equation A:

$$(w_1 \ w_2 \ w_3) = \lambda_1(0.5 \ 0.25 \ 0.125) + \lambda_2(0.4 \ 0.15 \ 0.225) - \lambda_3(0.3 \ 0.75 \ 0.325) - \lambda_4(0.2 \ 0.65 \ 0.425)$$

$$\lambda_1 = 4.5, \lambda_2 = 0, \lambda_3 = 1.5, \lambda_4 = 0$$

$$w_1 = 4.5*0.5 - 1.5*0.3 = 1.8$$

$$w_2 = 4.5*0.25 - 1.5*0.75 = 0$$

$$w_3 = 4.5*0.125 - 1.5*0.325 = 0.075$$

$$\mathbf{w} = 1.8, \mathbf{w}_2=0, \mathbf{w}_3 = 0.075$$

$$\frac{\partial L}{\partial w_b} = \sum_{k=1}^K \lambda_k y_k = 0$$

Equation B:

$$\lambda_1*1 + \lambda_2*1 + \lambda_3*(-1) + \lambda_4*(-1) \neq 0 \text{ (our } \lambda \text{ not satisfy KKT conditions)}$$

Equation C:

$$\lambda_k > 0 \text{ (satisfy KKT conditions)}$$

$$\text{Equation D (support gutters): } \mathbf{w}^T \mathbf{x}_k + w_b = \pm 1.$$

$$\lambda_1 [+1 ((w_1 \ w_2 \ w_3) (0.5 \ 0.25 \ 0.125) + w_b) - 1] = 0$$

$$\lambda_2 [+1 ((w_1 \ w_2 \ w_3) (0.4 \ 0.15 \ 0.225) + w_b) - 1] = 0$$

$$\lambda_3 [-1 ((w_1 \ w_2 \ w_3) (0.3 \ 0.75 \ 0.325) + w_b) - 1] = 0$$

$$\lambda_4 [-1 ((w_1 \ w_2 \ w_3) (0.2 \ 0.65 \ 0.425) + w_b) - 1] = 0$$

Hence: prediction $\mathbf{w}^T \mathbf{x} + w_b$

$$\mathbf{w} = 1.8, \mathbf{w}_2=0, \mathbf{w}_3 = 0.075$$

$$\mathbf{w}^T \mathbf{x} + w_b = 1.8*1 + 0*1 + 0.075*0 + 1 = 2.8 > 0 \text{ (Class 1)}$$

2. Gutters from Equation D

$$w^T x_k + w_b = \pm 1$$

$$(w_1 \ w_2 \ w_3)^T x_k + 1 = 1$$

$$(w_1 \ w_2 \ w_3)^T x_k + 1 = -1$$

3. $w_1 = 1.8, w_2 = 0, w_3 = 0.075$

$$\text{prediction } w^T x + w_b =$$

$$\text{Sample 1: } (1.8 \ 0 \ 0.075)^T (0.5 \ 0.25 \ 0.125) + w_b = 0.909375 + 1 = 1.909375$$

$$\text{Sample 2: } (1.8 \ 0 \ 0.075)^T (0.4 \ 0.15 \ 0.225) + w_b = 0.736875 + 1 = 1.736875$$

$$\text{Sample 3: } (1.8 \ 0 \ 0.075)^T (0.3 \ 0.75 \ 0.325) + w_b = 0.564375 + 1 = 1.564375$$

$$\text{Sample 4: } (1.8 \ 0 \ 0.075)^T (0.2 \ 0.65 \ 0.425) + w_b = 0.391875 + 1 = 1.391875$$

A sample falls into the margin if it is between the two gutters

$$-1 < w^T x + w_b < 1$$

All of sample above do not fall into the margin

4. prediction per SVM $w^T x + w_b$

If $w^T x + w_b > 0$ then x belong to Class 1.

If $w^T x + w_b < 0$ then x belong to Class 2.

The equation $w^T x + w_b = 0$ is the hyperplane.

If we use hard margin SVM, $w^T x + w_b > +1$ then x belong to Class 1.

Multiply the label y : $+1 * (w^T x + w_b) > +1$

But if it is misclassified, then: $+1 * (w^T x + w_b) < -1$

If an value < -1 , it will less than 0 and also 1

So condition **2,4,6** will hold and vice versa, 1,3,5 will hold if sample in the Class 2 misclassified with Class 1.