Machine Learning - 2020-2021
Universitat Pompeu Fabra

Problems 6A: Support Vector Machine

Goal

The goal of this practice is to understand how Support Vector Machine is defined and work in supervised classification problems.

Exercises

1. Consider two classes: class $C_1 = \{\mathbf{x}_1 = \left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)\}$ with a label 1 and class $C_2 = \{\mathbf{x}_2 = \left(\begin{smallmatrix}-1\\-2\end{smallmatrix}\right), \mathbf{x}_3 = \left(\begin{smallmatrix}-2\\-1\end{smallmatrix}\right)\}$ with label -1.

   (a) Plot the points and write in parametric form the Support Vector Machine (SVM) classifier $g(\mathbf{x})$.

   Solution: The classifier will be of the form $g(\left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right)) = w_1x_1 + w_2x_2 + b$.

   (b) Write the primal problem, i.e., the optimisation problem that provides the SVM classifier.

   Solution:

$$\min_{\mathbf{w}} \frac{1}{2}\|\mathbf{w}\|_2^2$$
$$\text{subject to } y^{(n)}\left(\mathbf{w}^{\mathrm{T}}\mathbf{x}^{(n)} + b\right) \geq 1 \qquad \forall n = 1, 2, 3$$

   where $\left\{\left(\mathbf{x}^{(n)}, y^{(n)}\right)\right\}$ are the training pairs.

   (c) Write the dual Lagrangian problem. Solve it and compute the values of α_i, $i = 1, 2, 3$.

   Solution: The dual Lagrangian is

$$\min_{\alpha} \mathcal{L}_D(\alpha) = (\alpha_1, \alpha_2, \alpha_3)\left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right) - \frac{1}{2}(\alpha_1, \alpha_2, \alpha_3)\left(\begin{smallmatrix}5&5&4\\5&5&4\\4&4&5\end{smallmatrix}\right)\left(\begin{smallmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{smallmatrix}\right)$$
$$\text{subject to } \alpha_n \geq 0 \ \ \forall n = 1, 2, 3 \text{ and } \sum_{n=1}^{3}\alpha_n y^{(n)} = 0$$

   To obtain the values of α_i, we first have to derive $\mathcal{L}_D$ with respect to α.

$$\frac{d}{d\alpha}\mathcal{L}_D(\alpha) = \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right) - \left(\begin{smallmatrix}5&5&4\\5&5&4\\4&4&5\end{smallmatrix}\right)\left(\begin{smallmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{smallmatrix}\right) = \mathbf{0}$$

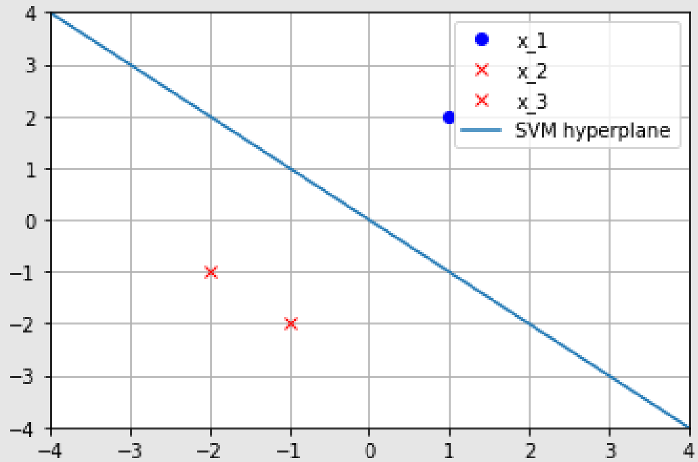
   This is a dependent system of equations, thus, we have to add the condition derived from the primal Lagrangian problem: $\sum_{i=1}^{3}\alpha_i y_i = \alpha_1 - \alpha_2 - \alpha_3 = 0$.
   Solving this system of equations we get $\alpha_2 = 0$, $\alpha_1 = \alpha_3 = \frac{1}{9}$.

   (d) Write the final classifier $g(\mathbf{x})$ and draw the decision hyperplane.

   Solution: From the primal Lagrangian:

$$\mathbf{w} = \alpha_1\mathbf{x}_1 - \alpha_2\mathbf{x}_2 - \alpha_3\mathbf{x}_3 = \frac{1}{9}\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right) - \frac{1}{9}\left(\begin{smallmatrix}-2\\-1\end{smallmatrix}\right) = \frac{1}{3}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right).$$

   Therefore, the classifier is be $g(\left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right)) = \frac{1}{3}x_1 + \frac{1}{3}x_2 + b$.
   Finally, imposing the two support vectors we get $b = 0$. The final classifier is: $g(\left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right)) = \frac{1}{3}x_1 + \frac{1}{3}x_2$.



(e) What is the margin value of the obtained classifier?

Solution: margin $= \dfrac{2}{||\mathbf{w}||} = \dfrac{2}{\frac{1}{3}\sqrt{2}} = 3\sqrt{2}$

2. Consider two classes: the class $C_1 = \{\mathbf{x}_1 = \binom{0}{-1}\}$ with a label 1 and the class $C_2 = \{\mathbf{x}_2 = \binom{0}{1}, \mathbf{x}_3 = \binom{1}{1}\}$ with label -1.

(a) Plot the points and write in parametric form the Support Vector Machine (SVM) classifier $g(\mathbf{x})$.

Solution: The classifier will be: $g(\binom{x_1}{x_2}) = w_1x_1 + w_2x_2 + b$

(b) Write the primal problem, i.e., the optimisation problem that provides the SVM classifier.

Solution: Same as exercise 1.

(c) Write the dual Lagrangian $\mathcal{L}_D$ and derive the values of $\alpha_i, i = 1, 2, 3$.

Solution: Similarly to Exercise 1:

$$\mathcal{L}_D\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix} = (\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{1}{2}(\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix}1&1&1\\1&1&1\\1&1&2\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix}$$

$$\frac{d}{d\alpha}\mathcal{L}_D(\alpha) = \begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}1&1&1\\1&1&1\\1&1&2\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix} = \mathbf{0}$$

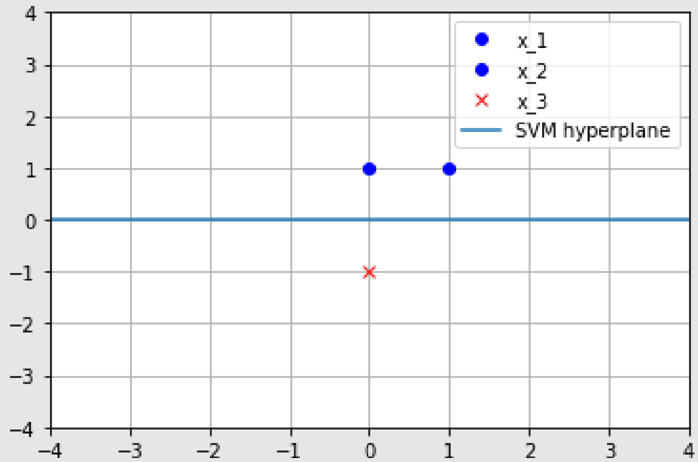
Recovering from the primal Lagrangian that $\alpha_1 - \alpha_2 - \alpha_3 = 0$ and solving the system of equations, we get $\alpha_3 = 0$, $\alpha_1 = \alpha_2 = \frac{1}{2}$.

(d) Write the final classifier $g(\mathbf{x})$ and draw the decision hyperplane.

Solution: Similarly to Exercise 1:

$$\mathbf{w} = \binom{0}{-1} \text{ and } b = 0.$$

Therefore, the classifier is $g(\binom{x_1}{x_2}) = -x_2$.



(e) What is the margin value of the obtained classifier?

Solution: margin $= \dfrac{2}{||\mathbf{w}||} = \dfrac{2}{1} = 2$