# SVM-Examples

① An SVM is trained with the following data.

$$
\begin{array}{cccc}
x_i & (-1,-1) & (1,1) & (0,2) \\
y_i & -1 & 1 & 1
\end{array}
$$

Note that there are 3 data instances here.

Therefore the gram matrix is $3 \times 3$.

Also there are 3 Lagrangian co-efficients $\alpha_1, \alpha_2, \alpha_3$

For polynomial kernel of degree 2, the gram matrix is

$$
\begin{bmatrix}
9 & \cancel{2}\ 1 & 1 \\
1 & 9 & 9 \\
1 & 9 & 25
\end{bmatrix}
\quad \left(\text{use formula } (x_i \cdot x_j + 1)^2\right)
$$

The optimization problem to be solved

$$
\alpha_1 + \alpha_2 + \alpha_3 - \frac{1}{2}\left[9\alpha_1^2 - 2\alpha_1\alpha_2 - 2\alpha_1\alpha_3 + 9\alpha_2^2 + 18\alpha_2\alpha_3 + 25\alpha_3^2\right]
$$

subject to

$$\alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_3 \geq 0, \quad -\alpha_1 + \alpha_2 + \alpha_3 = 0$$

Suppose an optimizer gave the values

$$\alpha_1 = 1/8 \quad \alpha_2 = 1/8 \quad \alpha_3 = 0$$

The support vectors are: $x_1, x_2$.

The classification boundary is given by

$$w'\phi(x) + b.$$

we can compute $b$ from any of the support vectors.
Say we use support vector $x_1$, $b = \frac{1}{y_1} - \sum_{j=1}^{k} \alpha_j y_j K(x_j, x_1)$

$$b = -1 - (1 * (-1) * 9 + 1 * 1 * 1)/8$$

$$= -1 - (-9 + 8/8)/8 = 0$$

To classify a new data point say $z = (-1, 0)$.

we compute $K(x_j, z)$ for all $x_j \in$ Support vectors.

here, $K(x_1, z) = 4 \quad K(x_2, z) = 0$.

$\therefore$ Classification rule using.

$$\sum_{i \in SV} \alpha_i K(x_i, x) + b$$

$$= -\frac{1}{8}(4) + \frac{1}{8}(0) < 0.$$

$\therefore$ Classified as $-1$

To Classify $z = (1, 0)$

$K(x_1, z) = 0 \quad K(x_2, z) = 4$

Classification =

$$-\frac{1}{8}(0) + \frac{1}{8}(4) > 0$$

$\therefore$ Classified as true.

# Adaboosting Example.

Consider a Single-dimension (1 feature) dataset

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $y$ | 1 | 1 | 1 | -1 | -1 | -1 | 1 | 1 | 1 | -1 |

Consider stumps of the form $x > v$ or $x < v$,

as the weak learner.

<u>Iteration 1</u>

| $p_0$ | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $p_6$ | $p_7$ | $p_8$ | $p_9$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |

best treshold between 2 & 3 to minimize error.

$$h_1(x) = I\,(x < 2.5)$$

$$\epsilon_1 = 0.3$$

$$\alpha_1 = 0.42$$

$q_i$ = 1.52 (for correct)

0.65 (for wrong)

normalized

| $p_0$ | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $p_6$ | $p_7$ | $p_8$ | $p_9$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.16 | 0.16 | 0.16 | ~~0.16~~ 0.07 |

$f_1(x) = 0.42\, I(x < 0.25)$

Iteration 2

Best treshold. between 8 & 9.

$h_2(x) = I(x < 8.5)$

$\epsilon_2 = 0.216$

$\alpha_2 = 0.64$

new probabilites (normalized)

| $p_0$ | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $p_6$ | $p_7$ | $p_8$ | $p_9$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.045 | 0.045 | 0.045 | 0.167 | 0.167 | 0.167 | 0.106 | 0.106 | 0.106 | 0.045 |

$f_2(x) = 0.42\, I(x < 0.25) + 0.64\, I(x < 8.5)$

$\vdots$

~~Conclusion~~