M339G: April 22nd, 2024.

We differentiate partially above w/ respect to $\beta_0, w_1, w_2, \dots, w_p$.

We get $\quad w_k = \sum_{i=1}^{n} \lambda_i y_i x_{ik} \quad$ for all $k = 1..p$

and $\quad \sum_{i=1}^{n} \lambda_i y_i = 0$

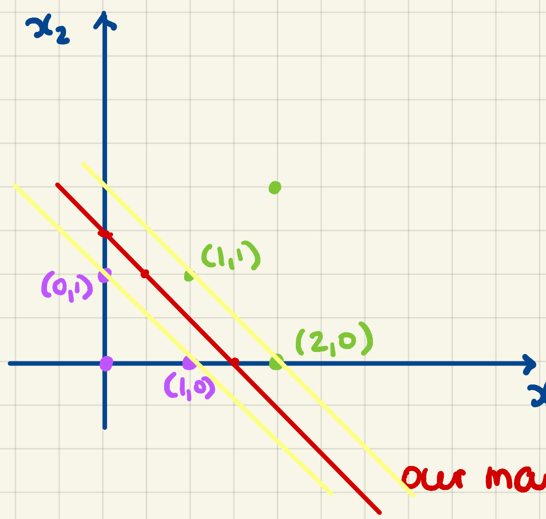
Moreover, by the KKT procedure, we have that

$$\lambda_i > 0 \quad \text{iff} \quad \boxed{y_i(\beta_0 + w_1 x_{i1} + w_2 x_{i2} + \dots + w_p x_{ip}) = 1} \leftarrow$$

i.e., the point x_i falls on the margin.

Problem. Consider these training data

	x_1	x_2	y
$i=1$	1	1	+1
$i=2$	2	2	+1
$i=3$	2	0	+1
$i=4$	0	0	−1
$i=5$	1	0	−1
$i=6$	0	1	−1



our main hyperplane candidate

w_1 and w_2 and $\beta_0 = ?$

$$\left.\begin{aligned} \beta_0 + w_1 \cdot 0 + w_2 \cdot 1 &= -1 \\ \beta_0 + w_1 \cdot 1 + w_2 \cdot 0 &= -1 \end{aligned}\right\} \quad w_1 = w_2 \qquad \beta_0 = -1 - w_1$$

$$\beta_0 + w_1 \cdot 1 + w_2 \cdot 1 = 1 \qquad -1 - \cancel{w_1} + \cancel{w_1} + w_1 = 1 \qquad w_1 = 2$$

$$\beta_0 + w_1 \cdot 2 + w_2 \cdot 0 = 1 \qquad\qquad w_2 = 2$$

$$\beta_0 = -3$$

$\Rightarrow$ <u>Eqn for the hyperplane</u> : $\underbrace{2}_{w_1} x_1 + \underbrace{2}_{w_2} x_2 = +3$

$$x_1 + x_2 = +\frac{3}{2}$$

$$\|w\|^2 = 2^2 + 2^2 = 8$$

$$\|w\| = 2\sqrt{2} \quad \Rightarrow \quad \frac{1}{M} = 2\sqrt{2} \quad \Rightarrow \quad M = \frac{1}{2\sqrt{2}} = \frac{\sqrt{2}}{4}$$

Task: Convince yourselves that the optimal margin does not increase if we discard (0,1) or (2,0).

Support Vector Classifier.

... is a relaxation of the maximal margin classifier.
It allows for a number of points to land on the wrong side of the margin or even the hyperplane. This is accomplished by introducing the slack ε_i for each point i.

Optimization problem.

$$\max_{\beta_0, \beta, \varepsilon} M$$

subject to $\sum_{j=1}^{p} \beta_j^2 = 1$

and $y_i(\beta_0 + \beta_1 x_{i1} + \cdots \beta_p x_{ip}) \geq M(1-\varepsilon_i)$ for all $i = 1 \ldots n$

with $\varepsilon_i \geq 0$ and $\sum_{i=1}^{n} \varepsilon_i \leq C$

slack variables

budget C

Q:

x	y
-1	$+1$
0	-1
1	$+1$