University of Toronto
CSC321, Winter 2018

HW 2: Xiangyu Kong 1002109620 kongxi16

1

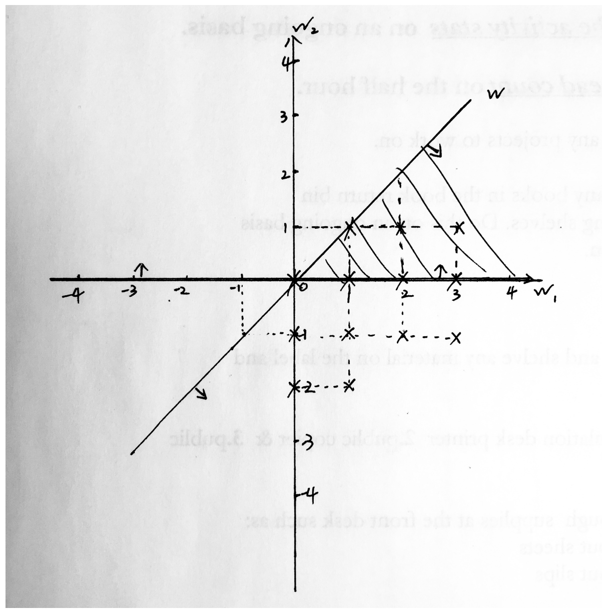

Figure 1: 1.1

2

1. $x^{(1)} = -1$ and $t^{(1)} = 1$ implies $w < 0$.
   However $x^{(3)} = 3, t^{(3)} = -1$ implies $w > 0$.
   Then this dataset is not convex, so it is not linearly separable.

2. For data $x^{(1)} = -1$, $y^{(1)} = \mathbf{w}^T \boldsymbol{\psi}(x^{(1)}) = -w_1 + w_2 > 0$.
   For data $x^{(2)} = 1$, $y^{(2)} = \mathbf{w}^T \boldsymbol{\psi}(x^{(2)}) = w_1 + w_2 < 0$.
   For data $x^{(3)} = 3$, $y^{(3)} = \mathbf{w}^T \boldsymbol{\psi}(x^{(3)}) = 3w_1 + 9w_2 > 0$
   A pair of values could be $w_1 = -2$ and $w_2 = 1$

3

$$
\begin{aligned}
\mathbf{y} &= \mathbf{X}\mathbf{w} + b\mathbf{1} \\
\frac{\partial \mathcal{E}}{\partial \mathbf{y}} &= \frac{1}{N}(sin(\mathbf{y} - \mathbf{t})) \\
\frac{\partial \mathcal{E}}{\partial \mathbf{w}} &= \frac{1}{N}\mathbf{X}^{\top} sin(\mathbf{y} - \mathbf{t}) \\
\frac{\partial \mathcal{E}}{\partial b} &= \frac{1}{N}(sin(\mathbf{y} - \mathbf{t}))
\end{aligned}
$$