Connectionist Neuron (Perceptron)

for $f = \text{sign}$

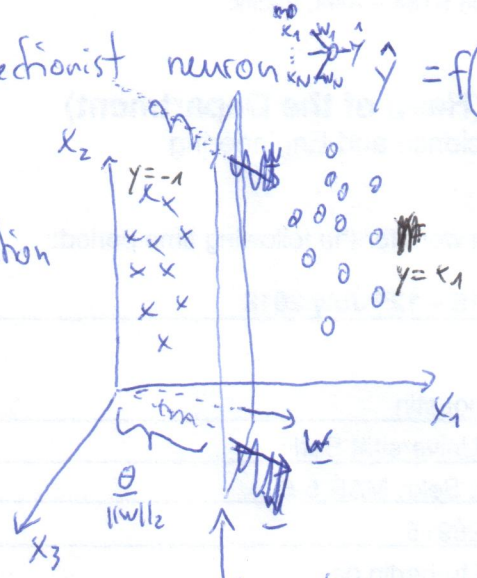
a) What does a connectionist neuron compute?

Connectionist neuron: $\hat{y} = f(\underline{w}^T \underline{x} - \Theta)$ with $\underline{w}, \underline{x} \in \mathbb{R}^N$, $\Theta \in \mathbb{R}$, $f: \mathbb{R} \to \mathbb{R}$

$\underline{x}$: data point; $\underline{w}, \Theta$: parameters; f: transfer function

$h := \underline{w}^T \underline{x} - \Theta$: extracted feature

example: classification



hyperplane $H = \{\underline{x} : \underline{w}^T \underline{x} = \Theta\}$

$$h \propto \frac{\underline{w}^T \underline{x}}{\|\underline{w}\|_2} - \frac{\Theta}{\|\underline{w}\|_2}$$

$\frac{\underline{w}^T \underline{x}}{\|\underline{w}\|_2}$: component of $\underline{x}$ in $\underline{w}$ direction

$\frac{\Theta}{\|\underline{w}\|_2}$: shift of hyperplane from origin since for $\tilde{\underline{x}} \in H$ we have $\frac{\underline{w}^T \tilde{\underline{x}}}{\|\underline{w}\|_2} = \frac{\Theta}{\|\underline{w}\|_2}$

b) Which effect have the weights and bias, respectively?

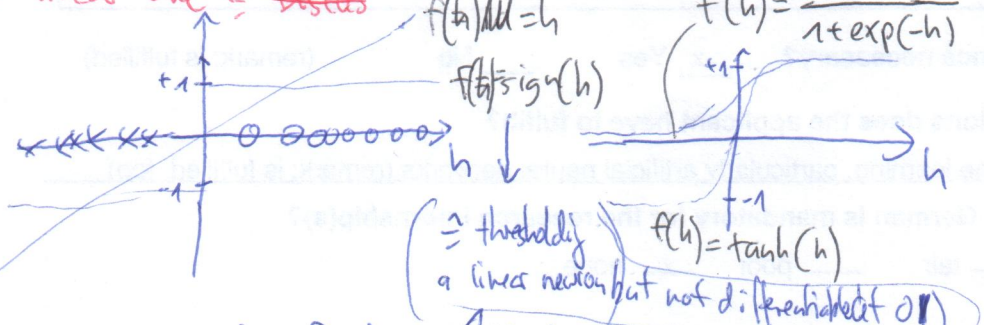
weights: orientation of hyperplane (normal vector $\underline{w}$)

bias: ~~distance~~ absolute position of hyperplane (in terms of the $\underline{w}$ direction)

(distance from zero in the $\underline{w}$ direction)

$\Rightarrow$ without bias could the above classification not be ~~performed~~ solved with a connectionist neuron

c) Why is a nonlinear transfer function beneficial to a linear one? ~~Discuss~~



$\pm$ threshold: a linear neuron but not differentiable at 0!

logistic ~~sigmoidal~~ is equivalent to hyperbolic tangent

$$\frac{1}{1+e^{-h}} = \frac{1}{2}\left[\tanh\left(\frac{h}{2}\right) + 1\right]$$

3 reasons:

i) classification ~~(not true)~~ $\rightarrow$ sign (tanh for continuous version)

ii) probabilities $\rightarrow$ logistic sigmoidal (yields values between 0 and 1)

iii) a multilayer perceptron reduces to a single layer for linear transfer function