Hebbian Learning — Learn Correlation in the data
Unsupervised Learning

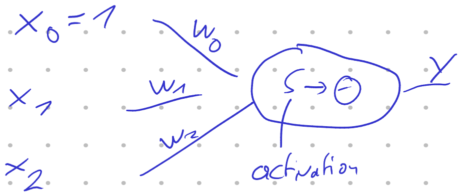
Perceptron — Learn Correlation between data x and labels y
Supervised Learning

Perceptron Learning Rule $\quad \Delta w_i = \varepsilon \, (\underset{\text{target}}{t} - \underset{\text{output}}{y}) \cdot x_i$

$$s = w_0 x_0 + w_1 x_1 + w_2 x_2$$

$$\underset{\text{output}}{y} = \begin{cases} 1 & \text{if } s > 0 \\ 0 & \text{else} \end{cases}$$

Logic Perceptron



And Problem

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

Incremental (Stochastic Gradient Descent)
threshold = 0, Learning rate = 2

x_0		1		1		1		1	1	1	1
x_1		1		0		1		1	0	1	0
x_2		1		0		0		0	1	1	0
s		5		-5		1		-3	·3	1	-7
y		1		0		1		0	0	1	0
t		1		0		0		0	0	1	0
w_0	-5		-5		-5		-7				
w_1	6		6		6		4				
w_2	4		4		4		4				

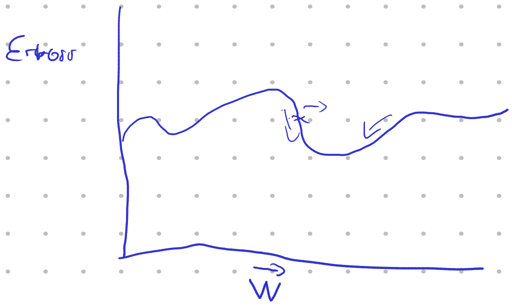
$$\Delta w_0 = 2(0-1) \cdot 1 = -2$$

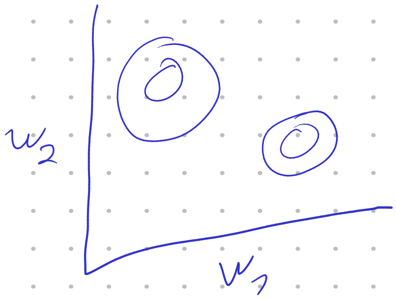
$$\Delta w_1 = 2(0-1) \cdot 1 = -2$$

$$\Delta w_2 = 2(0-1) \cdot 0 = 0$$

Optimization: Find the optimal weights to minimize the error function

Logic Perceptron: Possible to set weights by hand





Ideas to find optimal weights

- Random search
- Random local search
- Follow Gradient
    - Batch Gradient Descent
    - Mini Batch " "
    - Stochastic Gradient Descent

Gradient = Vector of partial derivatives ∇f

$f(a,b)$

$$\nabla f = \left[\frac{df}{da}, \frac{df}{db}\right]$$

Derivatives of simple functions

$f(a,b) = a \cdot b$ $\left[\frac{df}{da} = b \;;\; \frac{df}{db} = a\right]$ $y = wx$ $\frac{dy}{dw} = x$

$f(a,b) = a+b$ $[\,1\,,\,1\,]$ $y = b + wx$ $\frac{dy}{db} = 1$

$f(a,b) = \max(a,b)$ $\frac{df}{da} = \begin{cases} 1 & \text{if } a \geq b \\ 0 \end{cases}$ $\frac{df}{db} = \begin{cases} 1 & \text{if } b \geq a \\ 0 \end{cases}$

Chain Rule

$$\frac{df}{dx} = \frac{df}{dq} \cdot \frac{dq}{dx}$$

$f(x,y,z) = (x+y) \cdot z$

$f(q,z) = q \cdot z$ $q(x,y) = x+y$

$$\frac{df}{dx} = \frac{d(q \cdot z)}{dq} \cdot \frac{d(x+y)}{dx} = z \cdot 1 = z$$

Single Perceptron with Sigmoid Activation Function

$$y = \frac{1}{1+e^{-s}} = \frac{1}{1+e^{-1(w_0x_0 + w_1x_1 + w_2x_2)}}$$

$$\sigma(x) = \frac{1}{1+exp^{-x}}$$

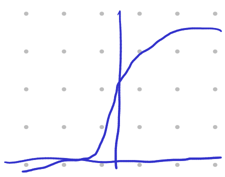
$f(x) = \frac{1}{x}$ $\frac{df}{dx} = -\frac{1}{x^2}$

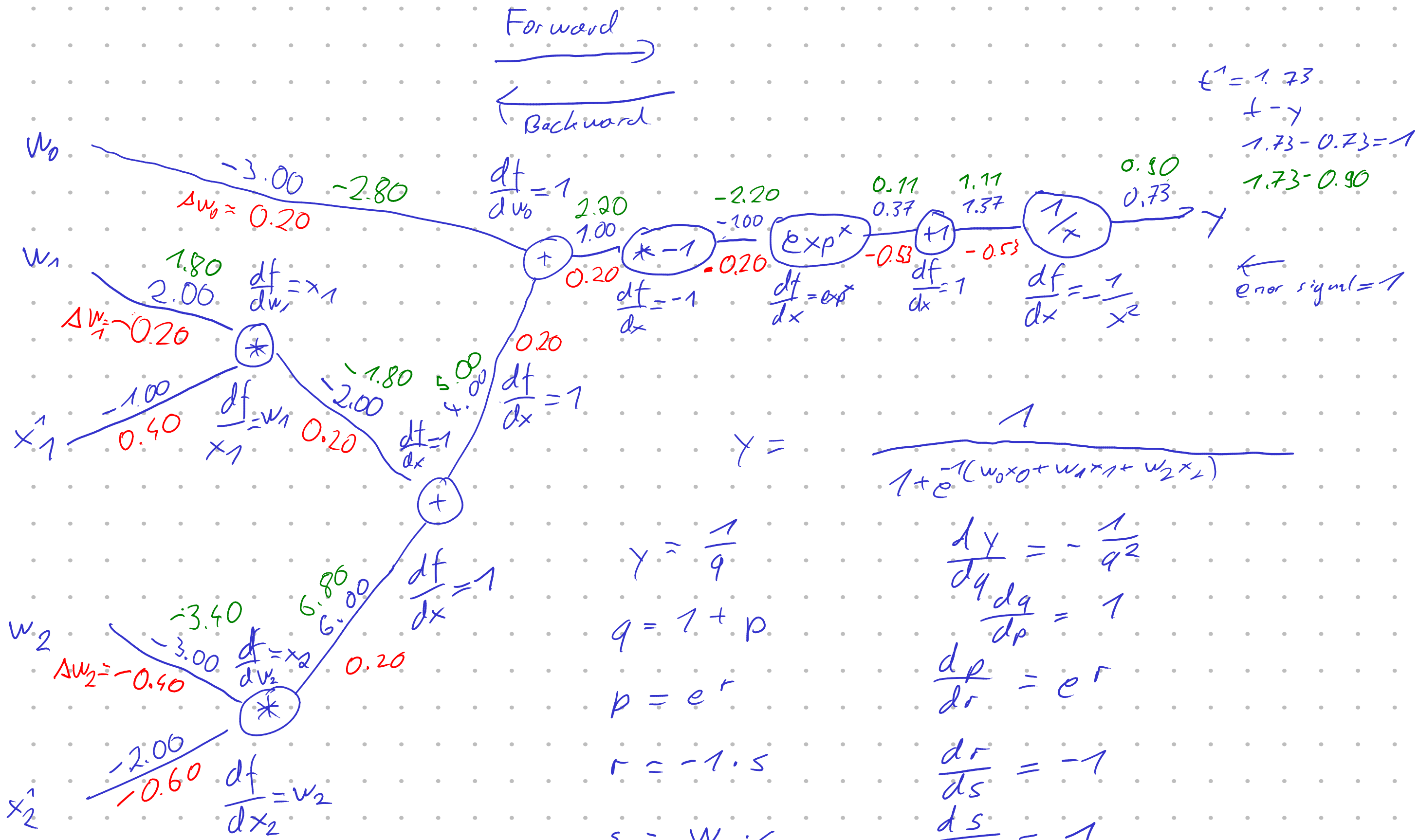
$f(x) = c+x$ $\frac{df}{dx} = 1$

$f(x) = e^x$ $\frac{df}{dx} = e^x$

$f(x) = ax$ $\frac{df}{dx} = a$

$$\sigma'(x) = (1-\sigma(x))(\sigma(x))$$





$$y = \frac{1}{1+e^{-1(w_0x_0+w_1x_1+w_2x_2)}}$$

$y = \frac{1}{q}$ $\qquad$ $\frac{dy}{dq} = -\frac{1}{q^2}$

$q = 1 + p$ $\qquad$ $\frac{dq}{dp} = 1$

$p = e^r$ $\qquad$ $\frac{dp}{dr} = e^r$

$r = -1 \cdot s$ $\qquad$ $\frac{dr}{ds} = -1$

$s = w_0 \cdot c$ $\qquad$ $\frac{ds}{dw_0} = 1$