Neural Networks
- Homework 2 -

Petr Lukin, Ivan Vishniakou, Evgeniya Ovchinnikova

Lecture date: 10 October 2016

1 Mind map

2 Exercises

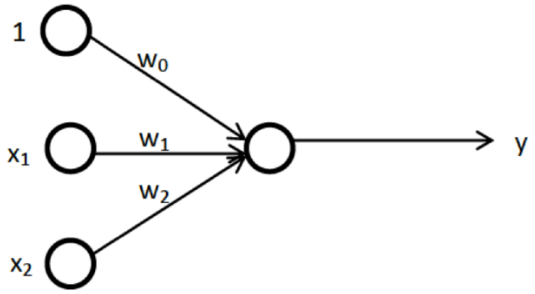
2.1 Exercise 1.13

2.2 Exercise 4

Derive the Delta rule for two ADALINE neural networks depicted in Fig. 1, Fig. 2.
Solution:

1 neural network.

Figure 1: Neural network 1.



Input: $\vec{x} = (x_0, x_1, x_2)$, where $x_0 = 1$
Weights: $\vec{w} = (w_0, w_1, w_2)$
Output: $y = \vec{x}\vec{w} = w_0 + x_1 w_1 + x_2 w_2$

$e = d - y = d - (w_0 + x_1 w_1 + x_2 w_2)$, where d is a desired value.

$E(w) = \frac{1}{2} e^2$

Each i^{th} weight with an should be moved in the direction $\frac{\partial E}{\partial w_i}$:

$\frac{\partial E}{\partial w_i} = x_i(d - (w_0 + x_1 w_1 + x_2 w_2))$

So, each step will change each weight by the following value:
$\Delta w_i = \eta x_i(d - (w_0 + x_1 w_1 + x_2 w_2))$,

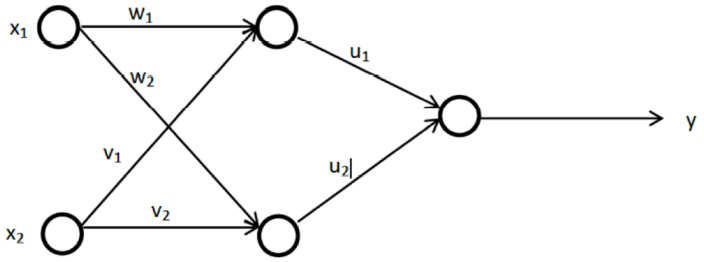
where η is a learning rate that should be chosen according to the specific of the certain case. It can be either a constant (0.5, 1, etc.) or variable, or adaptive.

Therefore the delta rule will take the following form:

$w_i^{new} = w_i^{old} + \eta x_i(d - (w_0 + x_1 w_1 + x_2 w_2))$

Neural network 2.

Figure 2: Neural network 2.



Input: $\vec{x} = (x_1, x_2)$
Weights: $\vec{w} = (w_1, w_2)$, $\vec{v} = (v_1, v_2)$, $\vec{u} = (u_1, u_2)$
Activation function is linear, so:
Output: $y = u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2)$

$e = d - y = d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$.

Each i^{th} weight with an should be moved in the direction $\frac{\partial E}{\partial weight_i}$:

$\frac{\partial E}{\partial v_i} = x_1 u_i(d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$

$\frac{\partial E}{\partial w_i} = x_2 u_i(d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$

$\frac{\partial E}{\partial u_i} = (x_1 w_i + x_2 v_i)(d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$

So, each step will change each weight by the following value:

$\Delta w_i = \eta_w x_1 u_i(d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$,
$\Delta v_i = \eta_v x_2 u_i(d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$
$\Delta u_i = \eta_u (x_1 w_i + x_2 v_i)(d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$

Therefore the delta rule will take the following form:

$w_i^{new} = w_i^{old} + \eta x_i(d - (w_0 + x_1 w_1 + x_2 w_2))$
$v_i^{new} = v_i^{old} + \eta_v x_2 u_i(d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$
$u_i^{new} = u_i^{old} + \eta_u (x_1 w_i + x_2 v_i)(d - (u_1(x_1 w_1 + x_2 v_1) + u_2(x_1 w_2 + x_2 v_2))$