2011, 5, 19

Research Note No. 907

Back Propagation in Universal Learning Networks (ULNs)

1. Introduction

The research note No. 907 reviews the back propagation algorithm of Universal Learning Networks, which is described in the paper titled Universal Learning Network and its Application to Chaos Control published in Neural Networks, vol 13, pp. 239–253, 2000.

2. Back Propagation Algorithm

The backpropagation algorithm of ULNs is described as follows (see Fig 1), which is essentially the same as the conventional backpropagation algorithm.

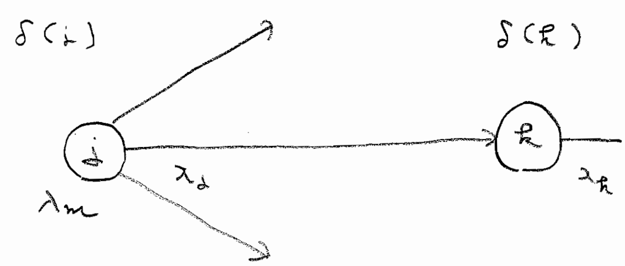

Fig. 1 Back Propagation of δ

$$\lambda_m \leftarrow \lambda_m - \alpha \frac{\partial^+ E}{\partial \lambda_m}$$

$$\frac{\partial^+ E}{\partial \lambda_m} = \sum_{n \in N} \left[\frac{\partial^+ E}{\partial x_n} \frac{\partial x_n}{\partial \lambda_m} \right] + \frac{\partial E}{\partial \lambda_m}$$

$$\delta(i) = \frac{\partial^+ E}{\partial x_i}$$

$$\delta(i) = \sum_{k \in A(i)} \left[\frac{\partial x_k}{\partial x_i} \delta(k) \right] + \frac{\partial E}{\partial x_i}$$

where

λ_m : the mth parameter

α : learning coefficient

$\frac{\partial^+ E}{\partial \lambda_m}$: ordered derivative of E with respect to λ_m

$\frac{\partial^+ E}{\partial x_i}$ ordered derivative of E with respect to x_i

E : evaluation function

N : set of suffixes of nodes in ULNs

x_n : the nth node in ULNs

$\frac{\partial x_n}{\partial \lambda_m}$, $\frac{\partial E}{\partial \lambda_m}$, $\frac{\partial x_k}{\partial x_i}$, $\frac{\partial E}{\partial x_i}$: partial derivative

$A(i)$: set of suffixes of nodes connected from x_i

3. An Example of Calculating Back propagation Algorithm

In this section, how to execute the back propagation algorithm is explained using an example shown in Fig. 2

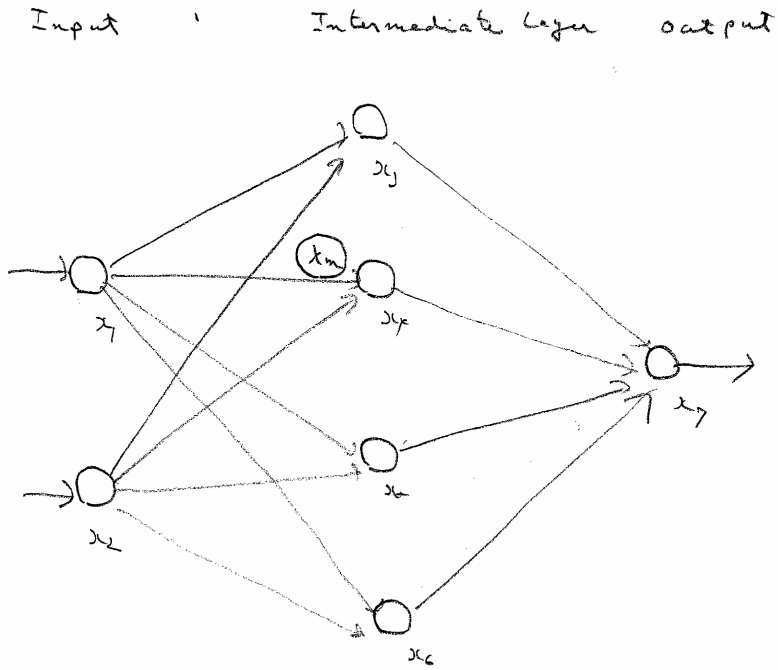

Fig. 2 An example of ULNs

Let us suppose the following evaluation function.

$$E = \sum_{l \in L} \left(\bar{x}_7(l) - x_7(l) \right)^2$$

where

$x_7(l)$: output of output node x_7 for the lth input

$\bar{x}_7(l)$: target output of output node x_7 for the lth input

L : set of suffixes of inputs

And also we suppose that parameter λ_m belongs to node x_4,

then λ_m is adjusted as follows.

$$\lambda_m \leftarrow \lambda_m - \alpha \frac{\partial^+ E}{\partial \lambda_m}$$

$$\frac{\partial^+ E}{\partial \lambda_m} = \delta(4) \frac{\partial x_4}{\partial \lambda_m}$$

(• $\frac{\partial x_n}{\partial \lambda_m} = 0$ except $\frac{\partial x_4}{\partial \lambda_m}$

• $\frac{\partial E}{\partial \lambda_m} = 0$)

$$\delta(4) = \frac{\partial x_7}{\partial x_4} \delta(7)$$

(• output of x_4 is connected to only x_7

• $\frac{\partial E}{\partial x_4} = 0$)

$$\delta(7) = \frac{\partial E}{\partial x_7} = -2 \sum_{e \in L} \left(\bar{x}_7(e) - x_7(e) \right)$$

(• the first term of δ equation is zero, because there is no node from x_7

• only $\frac{\partial E}{\partial x_7}$ is calculated)