DEFANET2 - ADVANCEMENTS OF A DETERMINISTIC FUNCTION APPROXIMATOR

Wolfgang J. Daunicht

Neurologisches Therapiecentrum, Hohensandweg 37, D-4000 Düsseldorf, Germany

The approximation of an arbitrary continuous function by a neural network is, although possible in principle, an uncertain and time-consuming undertaking. First of all, a training procedure usually requires an unknown and generally high number of steps. Second, it is difficult to assess a size of a network, that is capable to learn the desired function; therefore it may be necessary to iterate the learning sequence to find a network of suitable size. Third, even if the network is capable of approximating the desired function, the learning procedure – depending on the initialization – may end up in a local minimum far from the global optimum. Again, the solution may be to repeat the training with different initializations.

One possible solution to such problems might be to remove part of the uncertainties in the topology of the network and in the values of the synaptic weights, so that the learning procedure is characterized by only one global minimum and it becomes possible to calculate the synaptic weights using a finite algorithm. Such a concept using a special three-layered architecture called DEFAnet has been suggested and it has been proven, that a DEFAnet is in fact a universal approximator [1].

In this advanced version DEFAnet2 a fast algorithm is presented, that allows to determine the synaptic weights so that the network function g is differentiable and matches the desired function f at all inputs $\underline{x}$ on the grid intersections indexed with the vector index $\underline{\lambda}$.

The basic idea to reduce the computational burden is to transform the problem into simpler sub-problems. In DEFAnet2 the problem is broken down in two steps. First, it is transformed into the inversion of a set of 2^n matrices W, then the inversion of these matrices is reduced to the inversion of n (by orders of magnitude smaller) matrices W^*, n being the input dimension. Each matrix W is a multiple Kronecker product (denoted by $\otimes$) of some subset of the matrices W^*. The inverse of such a matrix W can be composed of the inverses of the factor matrices W_i by Kronecker multiplication

$$W^{-1} = \bigotimes_{i=1}^{j} (W_i^*)^{-1} \tag{1}$$

This formula assures considerable computational load reduction, even if the factor matrices have to be solved by means of a general inversion algorithm. However, it turns out that the factor matrices W^* are not necessarily of a general type, instead they can be made easier to invert, when mild restrictions are obeyed. For example, when all activation functions of those second first layer neurons receiving input from the same i-th neuron in the zeroth layer are identical, then the respective factor matrix W_i^* is a Toeplitz matrix. Toeplitz matrices, which are symmetrical and very regular, can be inverted quite effectively [2].

By holding to another restriction, i.e. using sigmoidal activation functions in the first layer that saturate exactly at 0 and 1, instead of approaching these values asymptotically, the factor matrices become symmetrical 'ribbon'-shaped, i.e. the elements are zero except for a narrow ribbon along the diagonal. Such matrices can be inverted even more effectively by an algorithm based on Cramer's rule.

Thus DEFAnet2 provides precise differentiable function approximator networks with efficient algorithms for weight calculation.

References

[1] W.J. Daunicht. DEFAnet - a deterministic neural network concept for function approximation. *Neural Networks*, 4:839–845, 1991.

[2] E.A. Robinson and S. Treitel. The Toeplitz recursion. In *Geophysical signal analysis*, pages 163–169. Englewood Cliffs, N.J., Prentice Hall, 1980.

Acknowledgements: This work was supported by DFG grant Da 199/2-1.

S. Gielen et al. (eds.), *ICANN '93*