TOPOLOGICAL LEARNING OF KERNEL NEURAL NETWORKS

A. Dorogov and A. Alekseev

St.-Petersburg State Electrotechnical University «LETI», prof. Popova, 5, RUSSIA, dorv@actor.ru.

ABSTRACT

For neural networks with kernel organization the method of topology construction based on data structural characteristics is offered. Algorithm of topological learning is shown on an example of kernel neural network with a regular structure. The work is maintained by the grant of РФФИ 00-01-00670a.

1. INTRODUCTION

Neural networks with kernel organization [1] are a structured variant of the multilayer perceptrons. Because of connection restrictions in kernel neural networks it is possible to receive a high velocity of data processing, and possibilities of broad variation of its structure allow to ensure adaptation networks to a structure of data. The necessity in similar networks arises in systems of real time for want of large volumes of a treated information. On fig. 1 a regular structural model of kernel neural network [2] is shown.

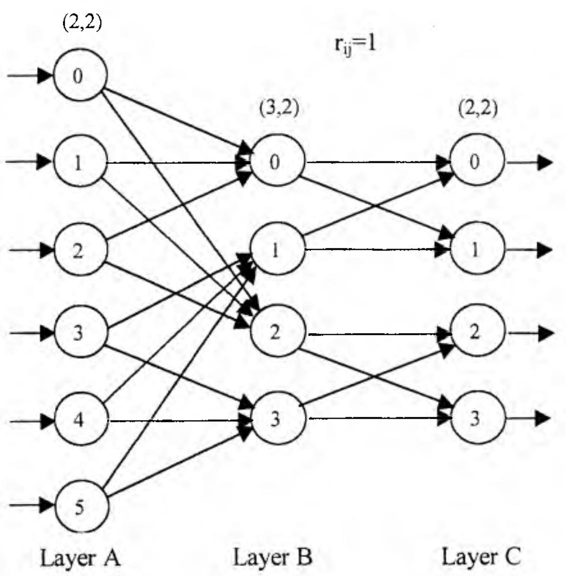

Figure 1. Structural model of neural network

On the structural model graph each top corresponds to a neural kernel, which can be considered as a single-layer perceptron of small dimensionality.

In the given example all neural kernels within of a layer are structural similar. Structure performance of a neural kernel is set by a pair of numbers *(p,g)*, where *p*- is dimensionality of receptive field of the kernel, and *g*- is number of its neurons (axons). It is postulated that neural kernel receptive fields are not intersected, and interkernel connections are injective [1].

Weight r_{ij} of an arc is equal to rank of the interkernel connection operator. For the given example all arcs have of weight equaled to unit. On fig. 2 the topological image of the neural network is shown as a set of topological matrixes. Each matrix represents the template of a layer synaptic card, where the nonzero elements are marked by a numeral «1». In the topological matrix a line number corresponds to a receptor number, and a column number corresponds to a neuron number. The given image unequivocally corresponds to known neural network graphic image (where tops are formal neurons), but it is more compact. For the same structural model it is possible to construct a set of various topologies, which we shall name as topological realizations.

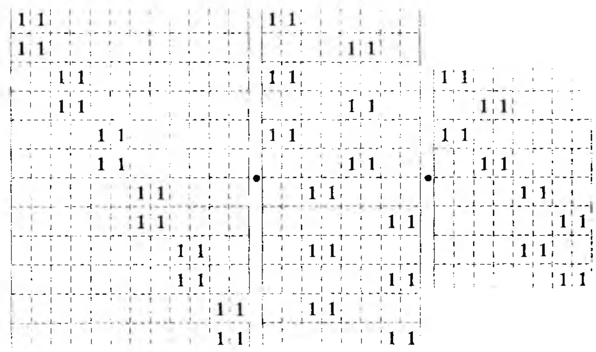

Figure 2. Topology realization of neural network

If the choice of topological realization is carried out because of learning data, than this operation can be treated as topological learning of neural networks.

2. ALGORITHM OF TOPOLOGY LEARNING

A neural network topology can be constructed by means of setting linear orders for receptors and axons in each network layer. For example, in neural layer A it is possible to assign a serial number for each receptor $U^A \in (0,1,2,K,P\text{-}1)$, and for each axon a serial number $V^A \in (0,1,2,K,G\text{-}1)$, where P,G are accordingly number of receptors and number of neurons in the layer. In turn for each neural kernel A_i with number of receptors equals p and number of neurons equals g it is possible to define local serial numbers

$$u_{A_i} \in (0,1,2,K,p\text{-}1) \text{ and } v_{A_i} \in (0,1,2,K,g\text{-}1)$$

Next pair of numerical maps

$$\sigma_{A_i}: U^A \rightarrow u_{A_i} \text{ and } \mu_{A_i}: V^A \rightarrow v_{A_i},$$

is named as neural kernel topology and determinated one to one correspondence between global and local serial numbers. It is convenient these maps to represent by following tables:

$$\sigma_{A_i} = \begin{pmatrix} U_0^A & U_1^A & \dots & U_{p-1}^A \\ 0 & 1 & \dots & p-1 \end{pmatrix},$$

$$\mu_{A_i} = \begin{pmatrix} V_0^A & V_1^A & \dots & V_{g-1}^A \\ 0 & 1 & \dots & g-1 \end{pmatrix}$$

In kernel networks receptive fields are not intersected, therefore it will be the following:

$$\sigma_{A_i} \cap \sigma_{A_j} = 0, \; \mu_{A_i} \cap \mu_{A_j} = 0$$

for $i \neq j$. An interlayer transition installs mutual dependence between axon topology of a preceding layer and receptor topology of a next layer, the algorithm of construction of dependent topology is indicated in articles [2,3].

Topological learning is connected only with a choice of terminal fields topology of a neural network (topology of input receptive field and topology of output axon field), the topology of the latent layers does not influence on the algorithm of data processing and can be selected to a certain extent arbitrary.

Structure of a neural network sets a hierarchy of the equivalent relations in terminal fields. Let's consider a principle of their formation.

Let B is a top (kernel) of some layer of a structural model. Let's name as afferent of the kernel B (designated further as *Afr*(B)) subset of input layer neural kernels connected by arcs with the top B. In a similar way subset of neural kernels of final layer connected by arcs with the top B, let's name as efferent of the top and we shall designate *Efr*(B). We shall name afferent and efferent also as terminal projections of the top. Let's limit our research to dense structures, for which the terminal projections of tops of identical neural layer either are not intersected or coincide. Thus, the terminal projections will derivate to partition of sets of terminal layers tops in not intersected classes. Tabl. 1 represents example of the terminal projections for structural model shown on fig.1.

Table 1. Terminal projections for layer tops

Layer	Afferents	Efferents
A	$\{A_0\}\{A_1\}\{A_2\}$ $\{A_3\}\{A_4\}\{A_5\}$	$\{C_0C_1C_2C_3\}$
B	$\{A_0A_1A_2\}\{A_3A_4A_5\}$	$\{C_0C_1\}\{C_2C_3\}$
C	$\{A_0A_1A_2A_3A_4A_5\}$	$\{C_0\}\{C_1\}\{C_2\}\{C_3\}$

It is visible from the table, that the terminal projections of neural kernels for all layers are enclosed and derivate a hierarchy of equivalence.

The equivalence of tops of a structural model induce equivalence in topology of terminal layers. In tabl. 2 there are the compositions of partitions of indication space, defining dimensionalities of equivalent topological classes.

Table 2. Terminal projections compositions

Layer	Afferent compositions	Efferent compositions
A	(2,2,2,2,2,2)	(8)
B	(6,6)	(4,4)
C	(12)	(2,2,2,2)

It is clear, that a neural network in a maximum degree will realize its potential to learning, when in each topological class includes independent indications. If some integrated criterion of indications independence is determined, then the purpose of topological learning it is possible to formulate as search of such topology of terminal layers for which the criterion achieves an extremum.

The integrated criterion can be constructed as the sum of valuations of a degree of indications independence on all classes of topological equivalence. Let's assume, that we know how to execute an estimation of the mutual indications independence (one from possible versions of such estimation is considered in article [4]). Let's show a principle of construction of algorithm of topological learning on next example.

Let learning set is given by the data table $[X,Y]$. Each table line corresponds to a learning example, and columns X and Y form indication space for an input and output of a neural network. The algorithm represents a step by step procedure, the number of steps is determined by depth of a hierarchical enclosure of topological equivalence. The afferent and efferent topologies are constructed independently. Let's for example the algorithm for afferent topology will be constructed.

From tabl. 1, 2 it follows, that the algorithm will consist of two steps. A composition of partitions for a layer B is equal (6,6), therefore with a specific degree of exactitude on the set of indications X we cover two groups of independent columns (we have six columns in each group). Let's assume, that the columns under numbers have been settled down as follows

$$T_0=\{0,5,1,4,2,3\} \quad T_1=\{9,8,6,7,10,11\}$$

From tabl. 1 it is visible, that the subsets of columns are covered by afferents:

$$\{A_0,A_1,A_2\} \qquad \{A_3,A_4,A_5\},$$

inducing topological equivalence with the composition (2,2,2). Therefore on the second step in each subset of indications T_0T_1 we cover three groups of independent indications (two indications in group). Let's assume, that the following partition is obtained:

$$T_0=\{\{0,1\},\{2,4\},\{3,5\}\},$$

$$T_1=\{\{9,8\},\{6,10\},\{7,11\}\}.$$

Comparing outcomes with numbers covering afferent sets, it is possible immediately to make out topological maps:

$$\sigma_0=\begin{pmatrix}0 & 1\\ 0 & 1\end{pmatrix},\quad \sigma_1=\begin{pmatrix}2 & 4\\ 0 & 1\end{pmatrix},$$

$$\sigma_2=\begin{pmatrix}3 & 5\\ 0 & 1\end{pmatrix}\quad \sigma_5=\begin{pmatrix}7 & 11\\ 0 & 1\end{pmatrix}$$

In a similar way algorithm is executed for indications space Y. Let's assume, that the columns under numbers have been settled down as follows

$$T_0=\{0,1,2,3\} \qquad T_1=\{4,5,6,7\}$$

for layer B and

$$T_0=\{\{0,3\},\{1,2\}\} \qquad T_1=\{\{5,6\},\{4,7\}\}$$

for layer C. Comparing outcomes with numbers covering efferent sets, it is possible immediately to make out topological maps:

$$\mu_0=\begin{pmatrix}0 & 3\\ 0 & 1\end{pmatrix},\quad \mu_1=\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix},$$

$$\mu_2=\begin{pmatrix}5 & 6\\ 0 & 1\end{pmatrix}\quad \mu_3=\begin{pmatrix}4 & 7\\ 0 & 1\end{pmatrix}.$$

The topology of the neural network after topology learning is shown on fig 3.

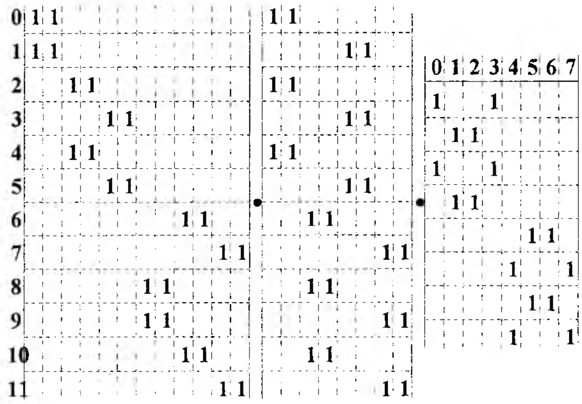

Figure 3. Realization of topology learned neural network

3. Conclusion

So as neural kernels in the afferents are not ordered, the plurality of possible topological tunings is supposed, anyone from them can be accepted as an outcome of topological learning. In deneral kernel networks the equivalence classes can be derivated as transitive closure on tolerance classes.

References

[1]Дорогов А.Ю, Алексеев А.А.. Категории ядерных нейронных сетей, Всерос. науч.-техн. конф. *Нейроинформатика-99*, Москва 20-22 января 1999г. Сб. науч. тр. Часть 1.-М.: 1999.-С.55-64.

[2]Дорогов А.Ю. Структурный синтез быстрых нейронных сетей, *Нейрокомпьютер*. **1** 1999.-С.11-24.

[3]Дорогов А.Ю.. Структурные модели и топологическое проектирование быстрых нейронных сетей, Докл. Междунар. Конф. "*Информационные средства и технологии*" 21-23 октября 1997г. г. Москва. / Т.1, М.: 1997, С.264-269.

[4] Dorogov A.Yu., Alekseev A.A. Information Criterion of Data Stucturization, Processing of *4-th International Conference "New Information Technologies"* (NITe'2000) Minsk, Belarus 5-7 December, 2000, vol. 1, p.12-14.