ハイパフォーマンス
コンピューティング 82−34
（２０００．８．５）

Symmetry on the contour map calculated by multi-layer neural networks

Hanxi ZHU, Tomoo AOYAMA*, and Umpei NAGASHIMA**
*The Faculty of Engineering, Miyazaki University
Gakuen Kibanadai-nishi 1-1, Miyazaki 889-2192, Japan
E-mail: aoyama_t@cc.miyazaki-u.ac.jp
**Computational Science, National Institute for Advanced Interdisciplinary Research
1-1-4 Higashi, Tsukuba, Ibaraki 305-8562, Japan
E-mail: umpei@nair.go.jp

Classifications by multi-layer neural networks have been applied in various fields, especially chemical industry and medical fields. It is important technologies that the compound activities are estimated before the synthesis. Since the relations among chemical-structure data and the activities are so non-linear, the estimation by neural networks is an effective tool. The standard deviations are complete zeroes at learning points at the observed data. However, there is a disputable point, because neural networks don't assure any values at non-learning points. The non-symmetric character in back-propagation learning is found, which makes a bias on the classifications. Some researchers would mention that the accurate classifications were obtained by the quantitative structure activity relationships; however, we imagine that characters of sigmoid function and appropriate sampling cause the results. We proposed a method in order to correct the bias, which based on the symmetry. The method introduced for the simple increasing functions, and extended to the general functions. We showed that the classifications corrected by the method were reasonable on several typical examples.

階層型ニューラルネットワークによって計算された等高線地図の対称性

朱含希、青山智夫*、長嶋雲兵**

*宮崎大学工学部電気電子工学科
**通商産業省工業技術院　産業技術融合領域研究所

階層型ニューラルネットワークを用いた分類法は種々の分野、特に化学工業、医薬分野に適用されている。化合物を合成する前にその活性を見積もることは重要な技術である。化学構造データと活性の間の相関は強い非線形である。ゆえにニューラルネットワークによる評価は効果的である。標準偏差は観測点、学習点では完全に零になる。しかし、そこに疑問をはさむ余地がある。なぜならニューラルネットワークは非学習点においてはどのような値も保証しない。バックプロパゲーション学習には非対称性が見いだされ、それは分類を偏らせる。ある研究者は定量的構造活性相関において正しい分類が得られた報告するかも知れない。我々はシグモイド関数と適切なサンプリングがその結果をもたらしたと思う。我々は分類偏りを対称性を基に補正する方法を提示する。

1. Introduction

Classifications by multi-layer neural networks are applied in various fields, which are medicines [1], agrichemicals [2], and catalysts.

It is important technologies that the compound's activities are estimated before synthesis in industry. Researchers have studied quantitative structure-activity relationships (QSAR) [3] that predict activities on use of structural observations of compounds. The relationships are so non-linear that the functions of neural networks are applicable to the estimations. Then, standard deviations are complete zeroes at learning points for the neural networks, and which are known compound's structural data. However, there is a disputable point, because neural networks don't assure any values at non-learning points. Therefore, it would be invalid expectations that QSAR by using neural networks gave accurate activities for unknown compounds. Some researchers would mention that the accurate estimations were obtained by the QSAR; however, we imagine that characters of sigmoid function and appropriate sampling cause the estimations.

2. Selections of neuron functions

2.1 Selection theory

Since any neuron-function is not assigned in the back-propagation learning, we can adopt other functions instead of the sigmoid function. The functions must have a differential possibility and extrapolating functions at unlearned points. If inappropriate functions for target phenomena are chosen, the estimations are unacceptable.

We adopt situations that target phenomenon is represented by continuous functions, and it is increasing plainly among observations, and it doesn't change rapidly. The restrictions are reasonable for QSAR. Under the restrictions, we considered selections of neuron functions on the second and third layers. The following pairs are practical:

(1) sigmoid for 2nd layer and sigmoid for 3rd layer,
(2) sigmoid and linear functions,
(3) sine and linear functions,
(4) quadratic and linear functions.

The case (1) is the basic and reference.

The case (2) is same as ANN that means "Analogue type Neural Network". Function of the ANN is,

$$f=\sum_{j}^{M} \mathbf{C}j\mathbf{S}j\left(\sum_{i}^{N} \mathbf{A}i\mathbf{x}i+\mathbf{B}\right) \qquad (1).$$

Where the number of 1st layer is N, that of 2nd layer is M, and 3rd layer is 1. If the number of 3rd layer is n, it rewrites f--->$\boldsymbol{f}$. The coefficient **C** is connected-value between neurons on 2nd and 3rd layer. **A** is connected-value between 1st and 2nd layer, and **B** is a bias. The **x** is input datum, and **S** means sigmoid function. Therefore, non-linear function of the ANN arises from sigmoid functions on 2nd layer only. If the sigmoid functions are replaced by others, the network-function is changed. As an example, if the replacement is ortho-normal set, the function of the ANN is equivalent to "otho-normal expansion". Since the equation (1) is finite expansion, the condition of the convergence is not zero but small positive number (about 10**-4). The convergence speed is faster with 30 times than that of plain neural network.

The method of ANN is near by radial function method. Recently genetic programming method [5] is well known as a superior prediction method, which generates explicit functions in order to represent phenomena. Cross terms between descriptors are also generated. So, one part in the variation space of the genetic programming method is equivalent to the ANN. The ANN has superior facility for extrapolations than that of all sigmoid functions. We believe the character is suitable in forecasting.

A selection reason for the case (3) and (4) is same as ANN. The (3) is derived from Fourier-series, and the (4) is done from Maclaurin expansion term in the sine function. They would be stiffer than that of sigmoid function; and the character is required in QSAR under uncertain observations for physiological activities. The convergence speeds are faster than that of ANN.

2.2 Contour maps for exclusive-OR problem

We investigated potential surfaces for the exclusive-OR problem in case of two-dimension. This is a simple non-linear classification problem, but the classification has some symmetrical operations. We are sure that the problem is suitable to study the potential surfaces calculated by various neural networks. The calculated surfaces are listed in figure 1-4.

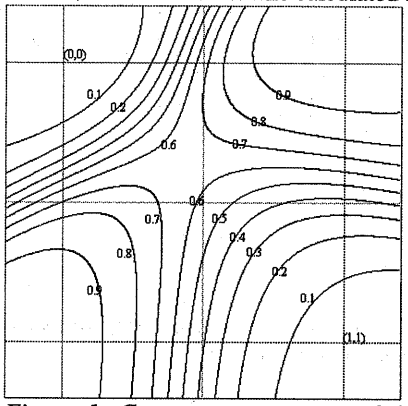

Figure 1. Contour map on use of {sigmoid, sigmoid} functions

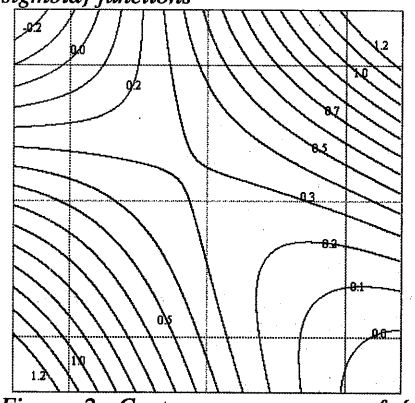

Figure 2. Contour map on use of {sigmoid, linear} functions, that is ANN.

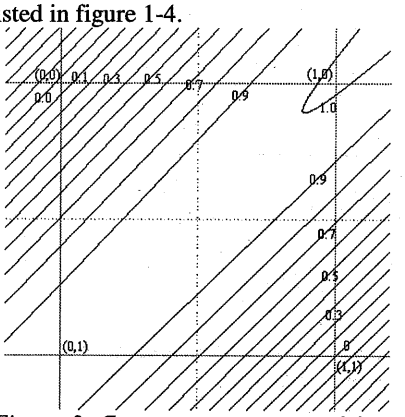

Figure 3. Contour map on use of {sine, linear} functions

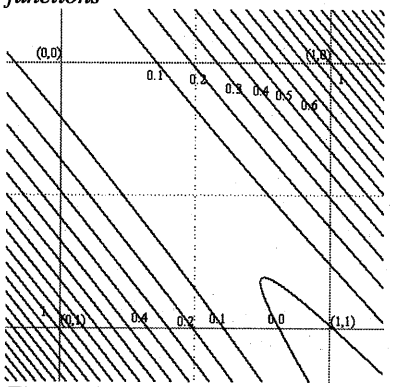

Figure 4. Contour map on use of {quadratic, linear} functions

Figures 1-4 indicate that the classifications are all true at learning points. But, there are some problems at unlearned points. That are symmetry, intervals of the contours, and network's output-values. The symmetry is not satisfied in all cases.

3. Introduction of symmetry

The symmetry is endowed with the data; it is not given by the learning, and the information processing never changes the symmetry. If non-symmetric responses were obtained from symmetric origins, they might be corrected by the symmetric operations.

The input data of two-dimensional exclusive-OR have following symmetry,

P(x--->1-x)D=1-D, P(y--->1-y)D=1-D, P(x--->1-x; y--->1-y)D=D,

where 0<x<1, 0<y<1, P() is transposing-operation, D is a logical value. We are sure that the relations must hold on outputs from neural networks. If following outputs were obtained,

O(0,0)=NN{x,y}, O(1,0)=NN{1-x,y}, O(0,1)=NN{x,1-y}, O(1,1)=NN{1-x,1-y},

where NN() means the processing in neural networks, the corrected output would be {O(0,0)+O(1,0)+O(0,1)+O(1,1)}/4. The correction is satisfied at any point in the two-dimensional

space. The results for the symmetric operations are listed in figure 5-8. Effects of the operations are remarkable, and especially, intervals of the contour map for ANN are regular, which is suitable character for the extrapolation.

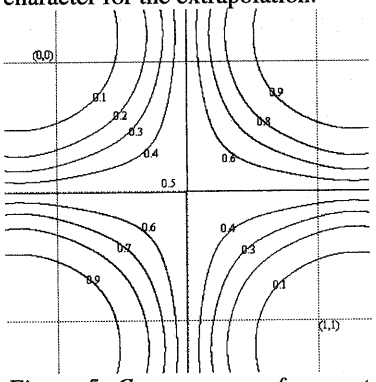

Figure 5. Contour map of symmetric output of {sigmoid, sigmoid} functions

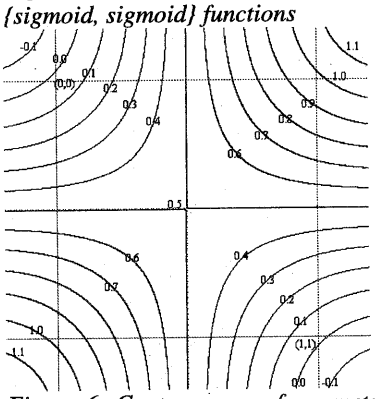

Figure 6. Contour map of symmetric output of {sigmoid, linear} functions

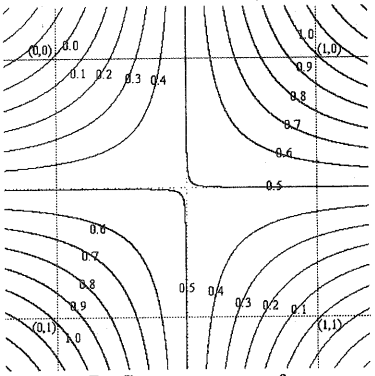

Figure 7. Contour map of symmetric output of {sine, linear} functions

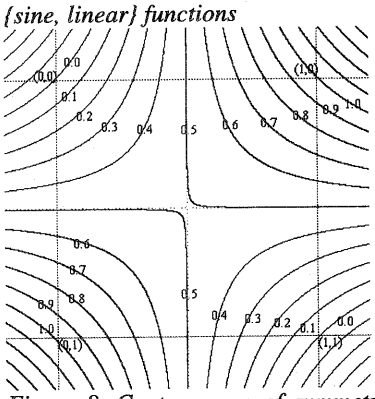

Figure 8. Contour map of symmetric output of {quadratic, linear} functions

4. Generalized symmetry

The observation data **X** are located in the space whose cardinal number is continuum. If we define a mapping that is **S**, then the operated data **X'=SX** are in another space. If the **X'** has a symmetry in the space, the results from the **X'** must have also same symmetric property. If not, the results should be corrected when the reverse mapping **~S(~SS=1)** can be defined. On the standpoint, we can make the classifications for QSAR by using {sigmoid, sigmoid} neural network, and checked the standpoint on a problem defined by table 1.

Table 1. Truth-value table

Locations	Values
(0,0)	0
(1,0)	1
(0,2/3)	1
(1,2/3)	0
(0,1)	0
(1,1)	1

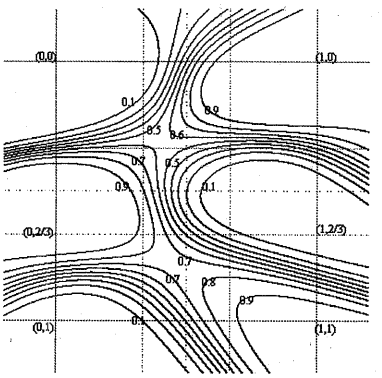

Figure 9. Original contour map calculated from Table 1

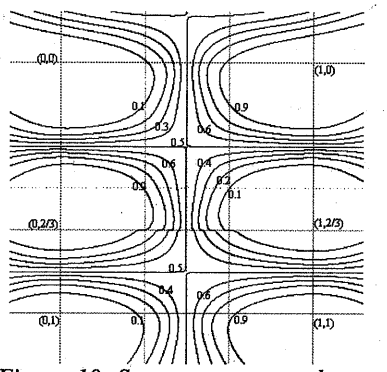

Figure 10. Symmetry corrected contour map of figure 9.

The figure 9 indicates that the calculated contour map has no symmetry, even if the learning data has symmetry (C2). The boundaries take curves around the learning points inappropriately. The confidence on extrapolations using the neural network would be low. Figure 10 is symmetry-corrected contour map that has complete C2-symmetry and reasonable symmetry in local zones. The boundaries take oval shapes around the learning points, that are reasonable. We accept the symmetry correction, and believe that it would extend the extrapolation ability of neural networks.

The symmetry correction is applicable in case of simple increasing functions. However, data for QSAR have much uncertainty, and then the correction method must be extended for general functions.

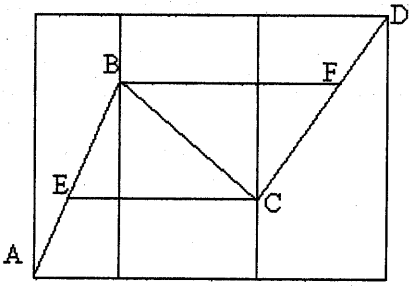

Figure 11. Simplified graph for an input data

There are many kinds of input data for QSAR. The most important of them is one-dimensional function that includes local maximum or minimum generally. It is increase and decrease of a function value in the figure 11 with the significant. Where a function is represented by the trace of A-B-C-D in the figure 11, and the maximum point is B, and the minimum point is C. The function relates with one representation that is same as the teaching data. The symmetry correction method makes the trace to be new trace {A-B, F-D}. The representation is not changed between B and F points. The new trace is plain increasing function; therefore, the correction can be defined. The representation between C and F points is same as that of E-B trace without scaling; and the representation between B and C points is reverse of that of E-B.

5. Conclusion

Classifications by multi-layer neural networks have been applied in various fields, especially chemical industry and medical fields. It is important technologies that the compound's activities are estimated before the synthesis. Since the relations among chemical-structure data and the activities are so non-linear, the estimation by neural networks is an effective tool. However, the non-symmetric character in back-propagation learning is found, it makes a bias on the estimation. We proposed a method in order to correct the bias, which based on the symmetry. The method introduced for the simple increasing functions, and extended to the general functions.

We showed that the classifications corrected by the method were reasonable on several typical examples.

Aknowledgement

The research was financially supported by the Sasagawa Scientific Research Grant from the Japan Science Society.

Reference

[1] Tomoo Aoyama, Yuji Suzuki, and Hiroshi Ichikawa,"Neural Network Applied to Structure-Activity Relationships", J. Medicinal Chemistry, 33. pp.905-908(1990).

Tomoo Aoyama, Yuji Suzuki, and Hiroshi Ichikawa,"Neural Networks Applied to Quantitative Structure-Activity Relationship Analysis", J. Medicinal Chemistry, 33. pp.2583-2590(1990).

[2] Tomoko Fukuda, Sumie Tajima, Hisato Saitoh, Umpei Nagashima, Haruo Hosoya, and Tomoo Aoyama,"Development of a neural network simulator for structure-activity correlation of molecules: Neco(5) Estimation of elution induce time and 80% elution time of polymer coated manure", submitted to J. Chemical Softwere.

[3] In the chemical or pharmacology fields, the most important problem has been the classification. We have always wished to develop new medicines that have useful properties and don't produce bad reactions. Here, we consider that medicines are synthesized by chemical reactions. In the chemical synthesizes, the medicines are constructed of many chemical functions and carriers. The functions are represented by the physical properties. Then, it might be concerned that properties for the medicines are predicted by using the physical properties. The prediction method is called as "quantitative structure-activity relationships (QSAR)". It has been the most important technique in drug manufacturing; and it is based on multi regression analysis. However, since the strong non-linearity is found among the functions, indeed, it is always found for medicines, the precision of the prediction is rather low.