

対称操作導入 BP 学習方程式の内挿外挿精度の改善手法

王謙軼, 渡邊隆由貴, 青山智夫*

宮崎大学工学部電気電子工学科

学習データに対称性が存在するとき、BP 学習方程式の結果にその対称性が反映される必要がある。しかし現実の BP 方程式の結果は一定区間の連続入力データに対し、前記必要性を満たさない。これを BP 方程式の「非対称性」という。データ処理によって被処理データ中に見いだされた本質が、そのデータから生成された近似関数であっても変化するのは望ましくない。この欠点を除去した対称 BP 学習方程式の導出は可能である。しかし過大な計算時間と解への収束性に問題が残る。その現実的解決策として BP 方程式の収束解に対し対称操作を施す。この解は対称 BP 学習方程式の解の一候補である。その解の内挿外挿精度を検討した。

An improvement technique for inter/extrapolate-precision, on use of symmetric operations in the BP learning

Qianyi WANG, Takayuki WATANABE, Tomoo AOYAMA*

The faculty of engineering, Miyazaki University
Gakuen Kibanadai-nishi 1-1, Miyazaki city, 889-2192, Japan.
t0b217u@cc.miyazaki-u.ac.jp

We propose a criterion in the neural-computing that is; when symmetrical property is found in the leaning data, the results, they may be a neural network, from the computing should also satisfy the property for all input data except the learning data. However, BP-learning equation does not satisfy the property. We call the fault as “unsymmetrical character of BP-learning.” It is inappropriate that the property in a subject is deformed by data processing. We eliminate the fault and can derive an equation; however, attainability towards one solution would not be assured because of many conditions that prevent the convergence-path in numerical calculation, and they bring excess CPU time. To avoid the trouble, we wish to adopt a correction method for the results from the BP-learning equation. The corrected result is a kind solution of the symmetrical BP-learning equation. We discussed the precision for inter/extrapolations of the corrected one.

Abstract

It is well known that multi-layer neural networks can be used as an excellent tool in the fields of chemistry and pharmacology. Our objective is to improve efficiency of the tool.

At first, we discussed the unsymmetrical property of BP-learning equation and derived a symmetrical learning-equation. However, the convergence for the equation would be less as probability and the equation requires excess CPU TIME; therefore, we propose the symmetrical correction method for responses of original BP-learning equation. We tried some typical model-calculations; as the results, we experienced efficiency of the correction was remarkable, and moreover the initial dependency for BP-learning was minimized.

Observed data are not symmetrical generally. We projected the data to a symmetrical space by using operators, and investigated expressions for the operators. Under the projection, the fitting ability near non-differential points is revised extremely, and high precise interpolation is enabled in the neural-computing.

Keywords: neural network, QSAR, symmetry, interpolation, extrapolation

1. Introduction

Classification functions of multi-layer neural networks have been applied in various fields. In chemical industry and medical fields, they have been important techniques; however in recent years, the intra/extrapolations are much more use. Especially, in the developments of medicines [1-6], agrochemicals [7,8], and catalysts [9,10], the intra/extrapolation techniques [11,12] are required now. Because estimations of the activities before compound synthesis are necessary to save the developing cost.

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 function parts and carriers. Their functions are represented by the physical properties. Then, it might be concerned that the properties for the medicines are predicted by using combinations of the physical properties. The prediction method is called quantitative structure-activity relationships (QSAR). It has been the most important technique in medicine manufacturing; and its theoretical base was the multi regression analysis [13]. However, since the non-linearity in the relations is always found, the prediction precision is rather low.

Neural networks has non-linear fitting functions, therefore, they may be useful tools. However, the multi-layer neural networks are considered originally as a classification tool, where information at the learning points [14] is dominant. While the objective of QSAR is intra/extrapolations, where information at learning points is not so important but that at unlearning intermediate points; therefore, some modifications are required. That are, selections of neuron functions and validity of calculated results at unlearning points should be considered [15]. On the considerations, the property of back propagation (BP) equation must be reconsidered [16].

2. Symmetrical BP-learning equation

When symmetric property is found in the observed data, the results after processing should have the property. BP-learning equation for multi-layer neural networks satisfy the requirement at the learning points; however, it does not satisfy at intermediate points among the learning ones [16]. We call the fault “unsymmetrical character” of BP-learning. It is inappropriate that property in the subject is deformed after information processing. We can derive an equation in which the fault is eliminated.

In original BP-learning equation, the energy “E” is defined as,

$$E = \sum_j (O_j - T_j)^2. \quad (\text{eq1})$$

Where O_j and T_j are output of j -th neuron on output-layer and expectation value for the j -th neuron respectively. The (eq1) is defined at the learning points. The differential is,

$$dE/dW = 0, \quad (\text{eq2})$$

where W is a connection-weight between neurons on different layers. BP-learning equation is derived from the (eq2).

When the symmetry is found in learning data, some elements of input data $\{X_i\}$ are equivalent each other. These elements are represented as $\{X_{ik}\}$, where index “ i ” means same irreducible representation category, even if index “ k ” is different. Then, (eq1) can be rewritten as one category,

$$E = \sum_k \{ \sum_j (O_j - T_j)^2 \}, \quad (k \text{ in “} i \text{”}), \quad (\text{eq3})$$

We wish to adopt the representation for the point group; however, it would be inappropriate slightly; because the response from the multi-layer neural networks has sometimes a kind of reverse symmetric relation, written as “ $1-X$ ”. When the reverse one is found, (eq3) should be rewritten,

$$E = \sum_k \{ \sum_j (O_j - T_j)^2 + R \sum_{j'} ((1 - O_{j'}) - T_{j'})^2 \}, \quad (\text{eq4})$$

where “ R ” is an operator, which enables $\sum$ -term only in case of the reverse relation.

We use an expression of the point group, including the reverse relation; i.e, as an example, we write C4-symmetry for exclusive-OR learning-data. Here, we denote that (eq1) is used only at the learning-points; however, (eq3 or 4) must be used at all points including intermediate regions. The difference causes excess CPU time and divergence of learning. The symmetrical BP-learning equation is a never calculated equation in practical. To avoid the problem, we wish to adopt a correction method for the results from original BP-learning equation, and wish to make the corrected one a solution of the symmetrical BP-learning equation.

3. Symmetrical correction

The symmetry is attached with data, i.e., C4-symmetry is found in exclusive OR problem, and C2 is found in linear increasing data. Therefore, if observation data have symmetrical property, outputs from the neural network learned by using the data should satisfy the symmetry. However, the results of BP learning have not such a character; therefore we often get non-symmetric responses from symmetric teaching data [15]. It is not acceptable for calculations of QSAR.

Since information processing makes no-changes to the symmetry, even if they are unlearning points in neural-computing, the break of symmetry is not acceptable. We have tried to introduce the symmetric property in BP learning; however, it was impossible. Therefore, we adopt a correction method for the results from neural networks that are learned completely; i.e., when a transformation

P() for input data operates, $P(X)=X'$, is found, we require the same responses for X' as well as X . We show an example.

The input data of two-dimensional exclusive-OR problem (C4) [15] have the following symmetry,
 $NN\{X,Y\}=D$, $NN\{P(X),Y\}=1-D$, $NN\{X,P(Y)\}=1-D$, $NN\{P(X),P(Y)\}=D$, (eq5)

$P(X)=1-X$, $P(Y)=1-Y$, (eq6)

where $0<X<1$, $0<Y<1$, $P()$ is a transforming-operator for input data, D is an output of the neural network, and $NN()$ means the processing in neural networks. We are sure the relations must hold, so we get a correction,

symmetry revised $D=\{NN(X,Y)+1-NN(1-X,Y)+1-NN(X,1-Y)+NN(1-X,1-Y)\}/4$. (eq7)

For C2-symmetry on the X-axis, we get,

symmetry revised $D=\{NN(X,Y)+1-NN(1-X,Y)\}/2$. (eq8)

The correction should be done at any point (X,Y) in the two-dimensional space. The results are listed in reference [15]. Effects of the operations are remarkable; we can suppress initial-guess independency of the networks, and especially intervals of the contour map for ANN are regular, which is suitable character for the QSAR intra/ extrapolations.

4. Generalized symmetry

Observation data are not symmetrical generally; the teaching data for BP-learning also are not, where we write them as X . We can't use the symmetrical correction for results of the X . However, since the correction revises inter/extrapolations functions of neural networks, we wish to extend the correction and to use them.

By a transformation operator S , we can get a relation $X'=SX$. If the X' is progression sequence, the symmetrical operations can be applicable. Thus, the problem is reduced to find the S operator. If the expression of the S is expressed as a discrete element's vector, it is a vector, $S=\{X'_0/X_0, X'_1/X_1, \dots\}$. The reverse transformation $\sim S$ is also defined from $\sim SS=1$; thus, $\sim S=\{X_0/X'_0, X_1/X'_1, \dots\}$.

Since we can seem X' is as sampling vector of linear function, we use the symmetrical correction for the responses of neural networks learned by using X' . The symmetrical responses have high precision inter/extrapolation ability, and they are projected by $\sim S$ operator, and give high precision inter/extrapolation as for the original X .

The method is only applicable at the learning points. For intermediate points X_m , $X_i < X_m < X_{i+1}$, we estimated S_m and $\sim S_m$ by using linear interpolations,

$$S_m = \{S_{i+1}(X_m - X_i) + S_i(X_{i+1} - X_m)\} / (X_{i+1} - X_i), \quad (\text{eq9})$$

$$\sim S_m = \{\sim S_{i+1}(\sim S_m - \sim S_i) + \sim S_i(\sim S_{i+1} - \sim S_m)\} / (\sim S_{i+1} - \sim S_i), \quad (\text{eq10})$$

The defect of (eq9/10) is unsymmetrical character for increasing/decreasing data.

Another expression is also derived as,

$$S_m = [\{X'_{i+1}(X'_m - X'_i) + X'_i(X'_{i+1} - X'_m)\} / (X'_{i+1} - X'_i)] / [\{X_{i+1}(X_m - X_i) + X_i(X_{i+1} - X_m)\} / (X_{i+1} - X_i)], \quad (\text{eq11})$$

$$\sim S_m = [\{X_{i+1}(X_m - X_i) + X_i(X_{i+1} - X_m)\} / (X_{i+1} - X_i)] / [\{X'_{i+1}(X'_m - X'_i) + X'_i(X'_{i+1} - X'_m)\} / (X'_{i+1} - X'_i)]. \quad (\text{eq12})$$

On use of (eq11/12), interpolations for symmetric neural networks can be executed in high precision; however, the extrapolation is not done. This is a limitation of the expression [17].

Since we calculated neural network on use of teaching data X and their input data Y , we can use

the response from the network as for Y_m , instead of X_m . We write it as $P(Y_m)$.

$$S_m = [\{X'_{i+1}(X'_m - X'_i) + X'_i(X'_{i+1} - X'_m)\} / (X'_{i+1} - X'_i)] / [\{P(Y_{i+1})(P(Y_m) - P(Y_i)) + P(Y_i)(P(Y_{i+1}) - P(Y_m))\} / (P(Y_{i+1}) - P(Y_i))], \quad (\text{eq13})$$

$$\sim S_m = [\{P(Y_{i+1})(P(Y_m) - P(Y_i)) + P(Y_i)(P(Y_{i+1}) - P(Y_m))\} / (P(Y_{i+1}) - P(Y_i))] / [\{X'_{i+1}(X'_m - X'_i) + X'_i(X'_{i+1} - X'_m)\} / (X'_{i+1} - X'_i)]. \quad (\text{eq14})$$

On case of the expression, extrapolation can be calculated, however the precision is lower.

5. Conclusion

Multi-layer neural networks have interpolating function originally, that is used for various application fields, i.e, estimations for relationships between chemical compounds and physiological activities. The neural networks have been a practical tool on the fields, and nowadays, reinforcement of the function is required continuously. Moreover recently, extrapolation is also required on industrial developments.

For the objectives, we considered some defects in multi-layer neural networks, and tried to eliminate them for symmetry of output-responses from neural networks, transformation on the learning (teaching) data, and introduced two techniques that are a combination for the symmetrical corrections for the responses and transformations from the ordinary space to symmetric one.

We tested effects of the techniques on many model of calculations [18]. As experience from them, we got useful techniques as for extrapolation, improvement of interpolations, and detecting a vertex in QSAR problems. We believe that the introduced techniques are practical, and give high performance calculations for QSAR.

6. Reference

- [1] T. Aoyama, Y. Suzuki, H. Ichikawa, "Neural Network Applied to Structure-Activity Relationships", J. Medicinal Chemistry, vol. 33, pp.905-908(1990).
- [2] T. Aoyama, Y. Suzuki, H. Ichikawa, "Neural Networks Applied to Quantitative Structure-Activity Relationship Analysis", J. Medicinal Chemistry, vol. 33, pp.2583-2590(1990).
- [3] H. Zhu, T. Aoyama, S. Tajima, T. Matsumoto, U. Nagashima, H. Hosoya, "Structure-Activity Correlation for Sweet/Bitter Classification by using Neural Networks", Proc. of AI in Sci. and Tech.(AISAT'2000), pp.74-79(2000).
- [4] G. Schneider, "Neural networks are useful tools for drug design", Neural Networks, vol. 13, pp.15-16(2000).
- [5] T. Borowski, M. Krol, E. Broclawik, T.C. Baranowski, L. Strekowski, M.J. Mokrosz, "Application of similarity matrices and genetic neural networks in qualitative structure-activity relationships of 2- or 4-(4-methylpiperazino)pyrimidines: 5-HT(2A)receptor antagonists, J. Medicinal Chemistry, vol. 43, pp. 1901-1909(2000).
- [6] F.R. Burden, D.A. Winker, "A quantitative structure-activity relationships model for the acute toxicity of substituted benzenes to *Tetrahymena pyriformis* using Bayesian-regularized neural networks", Chemical Research in Toxicology, vol. 13, pp.436-440(2000).
- [7] T.T. Bachmann, B. Leca, F. Vilatte, J.L. Marty, D. Fournier, R.D. Schmid, "Improved multianalyte

detection of organophosphates and carbamates with disposable multielectrode biosensors using recombinant mutants of *Drosophila* acetylcholinesterase and artificial neural networks”, *Biosensors and Bioelectronics*, vol. 15, pp. 193-201(2000).

[8] T. Fukuda, S. Tajima, H. Saitoh, U. Nagashima, H. Hosoya, T. 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 Software*.

[9] F. Larachi, “Neural network kinetic prediction of coke burn-off on spent $\text{MnO}_2/\text{CeO}_2$ wet oxidation catalysts, *Applied Catalysis B: Environmental*”, vol. 30, pp. 141-150(2000).

[10] H.C. Krijnsen, W.E.J. van Kooten, H.P.A. Calis, R.P. Verbeek, C.M. van den Bleek, “Evaluation of an artificial neural network for NO_x emission prediction from a transient diesel engine as a base for NO_x control”, *Canadian Journal of Chemical Engineering*, vol. 78, pp. 408-417(2000).

[11] T.I. Oprea, J. Gottfries, V. Sherbukhin, P. Svensson, T.C. Kühler, “Chemical information management in drug discovery: optimizing the computational and combinatorial chemistry interfaces”, *J. of Molecular Graphics and Modelling*, vol. 18, pp. 512-524(2000).

[12] H. Zhu, T. Aoyama, I. Yoshihara, S. Lee, W. Kim, “Structure-activity correlation relationships for chemical compounds precision indexes on use of neural networks”, *Proc. of 15th Korea Automatic Control Conference*, pp.481 (2000).

[13] I. Miriguchi, K. Komatsu, Y. Matsushita, *J. Medicinal Chemistry*, vol.23, pp.20(1979).

[14] The learning data are a set of vectors. Therefore, in a space whose dimension equals to the element’s number of the vector, the datum is a point. On the meanings, we call the data as learning points. As well as this manner, we use a word, “observed points.”

[15] H. Zhu, T. Aoyama, U. Nagashima, “Symmetry on the contour map calculated by multi-layer neural networks”, *Information Processing Society of Japan, SIG Notes*, 2000-HPC-82, pp.197-202.

[16] T. Aoyama, H. Zhu, U. Nagashima, “Quantitative structure activity relationships for medicine based on use of neural networks”, *Proc. of 15th Korea Automatic Control Conference*, pp.518 (2000), and CD-ROM.

[17] T. Aoyama, Q. Wang, U. Nagashima, “Reinforcement of extrapolation of multi-layer neural networks”, *Proc. of International Joint Conference of Neural Network 2001*(accepted).

[18] We wish to determine a testing function set for QSAR. Because, it is hard to discuss the approximation for QSAR theory by mathematical analysis; therefore, we can’t help using numerical calculations. Thus, we should select a reasonable testing set in reference [16].