

Supervised Learning Methods of Bilinear Neural Network Systems Using Discrete Data

Toshio Ito

Abstract—This paper presents supervised learning methods of neural networks called bilinear neural networks with time delay (BNN). A BNN system was proposed to analyze a weak nonlinear model. In this paper, we propose supervised learning methods of BNN systems using discrete data and continuous curves of the data obtained by curve fitting. We introduce a method for fitting a finite Fourier series to discrete data and show that the fitted curve can be created as the output from a BNN system. By using this fitting method, we propose a method for determining the optimal values for the coefficients of all connections for each neuron in BNN systems.

Index Terms—Supervised learning methods, neural networks, nonlinear model, discrete data, discrete Fourier transform, Fourier series, curve fitting.

I. INTRODUCTION

Artificial neural networks (NN) are a mathematical model inspired by biological neural networks (cf. [1]). Many researchers have studied NNs to solve artificial intelligence problems. NNs can solve various problems through learning such as supervised learning methods, unsupervised learning methods, or reinforcement learning methods. However, since nonlinearity is realized by applying strong nonlinear functions such as sigmoid functions for neuron models, it is hard to analyze the values of neurons for a given NN model.

On the other hand, a new type NN for analyzing a *weak nonlinear model* is proposed in [2]. In this paper, we refer to this as a bilinear neural network (BNN) system with time delay. A BNN is the model specialized for analyzing time series data, and has neurons, normal connections, and new connections, called *bilinear connections*, to realize weak nonlinearity. It is possible to represent simultaneous linear or nonlinear differential equations using BNNs. Moreover, as shown in [3], it is possible to analyze the output functions from all neurons in a given BNN system unlike the given NN model. As analysis for a weak nonlinear model, BNN systems are useful for human motion analysis. As one example of human motion, the walking motion is analyzed by using the acceleration data outputted from the accelerometer embedded in a mobile phone and BNN systems in [3]. BNN systems are currently adopted in the pedometers for humans; furthermore those are adopted in the pedometers for dogs [4].

When several physical data are given, it is essential to construct optimum differential equations, which express physical laws related to given physical data. In this paper, as a

method of constructing optimum differential equations, we propose new supervised learning methods of BNN systems using discrete data.

We need physical input and output data for the supervised learning methods. Many different types of sensors are available for extracting physical data. For example, in mobile phones, accelerometers are embedded. As for other digital sensors, there are gyro sensors, magnetic-field sensors, microphones, temperature sensors, pressure sensors, and so on. Although acceleration, angular velocity, geomagnetism, sound, temperature and pressure are typical physical concepts in continuous time, physical data generated by these sensors are discrete data. Moreover, we cannot usually know the original continuous-time functions that describe these physical data.

For applying discrete data to a continuous-time model, we use the curve-fitting in [5] to transform these discrete data into continuous-time functions. In the curve-fitting approach, the discrete data are fitted to a finite Fourier series. Moreover, the fitted curve can be obtained as an output from a BNN system. Therefore, the BNN system is used for the supervised learning methods that we propose.

Supervised learning methods of BNN systems using discrete data are proposed in [5]. However, the supervised learning methods in [5] are only applied to a part of all BNN systems which only have normal connections. As one of the contributions of this paper, we extend the supervised learning methods of a part of all BNN systems which only have normal connections to supervised learning methods of all BNN systems which have bilinear connections in addition to normal connections. Though this contribution, it becomes possible to apply the learning methods to all BNN systems, which have normal and bilinear connections, for the first time.

In this paper, we first propose supervised learning methods of BNN systems using discrete input and output data. As for one property of supervised learning methods that we propose, we prove that these supervised learning methods approximate supervised learning methods of BNN systems using continuous-time functions. We then consider supervised learning methods of BNN systems using discrete data with noise data.

II. CURVE FITTING OF DISCRETE DATA

A. Definition of the Fitted Curve for Discrete Data

We fix a positive integer n , assume that the total number of discrete data points is an odd number $2n+1$, and denote these discrete points by y_0, y_1, \dots, y_{2n} . Then, the discrete Fourier transform (DFT) (cf. [6]) is defined by

Manuscript received August 4, 2016; revised October 8, 2016.
T. Ito is with Fujitsu Laboratories Ltd., Kawasaki, 211-8588 Japan
(e-mail: ito.toshio@jp.fujitsu.com).

$$Y_q = \sum_{i=0}^{2n} y_i e^{\frac{-2\sqrt{-1}qi\pi}{2n+1}} \quad (q=0,1,\dots,2n), \quad (1)$$

and the inverse discrete Fourier transform (IDFT) is defined by

$$y_i = \frac{1}{2n+1} \sum_{q=0}^{2n} Y_q e^{\frac{2\sqrt{-1}qi\pi}{2n+1}} \quad (i=0,1,\dots,2n). \quad (2)$$

Suppose that y_0, y_1, \dots, y_{2n} are real numbers. By expressing each DFT component as $Y_q = a_q - \sqrt{-1}b_q$, and choosing the first $n+1$ components from all DFT components, we obtain the fitted curve of y_0, \dots, y_{2n} by

$$y(t) = \frac{a_0}{2n+1} + \frac{2}{2n+1} \sum_{q=1}^n a_q \cos\left(\frac{2q\pi t}{2n+1}\right) + \frac{2}{2n+1} \sum_{q=1}^n b_q \sin\left(\frac{2q\pi t}{2n+1}\right). \quad (3)$$

from [5]. When we define λ as $\lambda = 2\sqrt{-1}\pi / (2n+1)$, the fitted curve in (3) is equivalent to

$$y(t) = \frac{1}{2n+1} (Y_0 + \sum_{q=1}^n (Y_q e^{\lambda qt} + \bar{Y}_q e^{-\lambda qt})). \quad (4)$$

Moreover, when we compute Y_q in (1) for $q = -2n, \dots, -1$, we obtain $Y_q = \bar{Y}_{-q}$ and can change (4) to

$$y(t) = \frac{1}{2n+1} \sum_{q=-n}^n Y_q e^{\lambda qt}. \quad (5)$$

The function $y(t)$ is a finite Fourier series with period $2n+1$ and is therefore a *smooth* function whose derivatives in all orders exist and are continuous. Since each coefficient in $y(t)$ is a real number, $y(t)$ is also a real number if t is a real number.

The following lemma shows that the fitted curve defined above passes through all the discrete data points and that the curve fitting is the bijection from the set of discrete data to the set of finite Fourier series. Define the vector of discrete data as $\vec{y} = (y_0, y_1, \dots, y_{2n})$, and the fitted curve of \vec{y} as $f_{\vec{y}}(t)$. In addition, when a finite Fourier series with real coefficients

$$u(t) = c_0 + \sum_{q=1}^n (c_q \cos\left(\frac{2q\pi t}{2n+1}\right) + d_q \sin\left(\frac{2q\pi t}{2n+1}\right)) \quad (6)$$

is given, define the vector of discrete data obtained by sampling $u(t)$ as $\vec{u} = (u(0), \dots, u(2n))$. We then obtain the following lemma from [5]:

Lemma 1:

1. For any discrete data y_0, \dots, y_{2n} , the data satisfy $\vec{f}_{\vec{y}} = \vec{y}$, where $\vec{f}_{\vec{y}} = (f_{\vec{y}}(0), \dots, f_{\vec{y}}(2n))$.

2. Any finite Fourier series $u(t)$ in (6) satisfies $f_{\vec{u}}(t) = u(t)$.

B. BNN System to Output a Fitted Curve

We briefly introduce a BNN system (cf. [2]–[5]). When neurons in a BNN system are denoted by x_1, x_2, \dots, x_N , each neuron x_i has the following equation

$$\varepsilon_i \frac{dx_i}{dt} + x_i = \sum_j C_{ij} x_j + \sum_{k \leq l} C_{ikl} x_k x_l + g_i, \quad (7)$$

and outputs the function $x_i(t)$. The parameter ε_i is called the delay scale, and C_{ij} and C_{ikl} are constant coefficients of the connections. If C_{ij} is not zero, x_i has a normal connection, and if C_{ikl} is not zero, x_i has a bilinear connection. The function $g_i(t)$ is given by an external input. The model of (7) is shown in Fig. 1.

We introduce the BNN system to output the fitted curve proposed in [5]. Fig. 2 shows a BNN system for an oscillation. This BNN system outputs $x_1(t) = \cos(2q\pi t / (2n+1))$ with the initial value of $x_1(0) = 1$, and $x_2(t) = \sin(2q\pi t / (2n+1))$ with the initial value of $x_2(0) = 0$.

The BNN system for curve fitting is shown in Fig. 3. When discrete data y_0, \dots, y_{2n} are given, this system computes the values $c_0 = a_0 / (2n+1)$, $c_q = 2a_q / (2n+1)$ and $d_q = 2b_q / (2n+1)$ using (1), saves the values and uses them as the coefficients of the connections in the system. This system has n BNN systems for the oscillation. The functions produced from these BNN systems for the oscillation are multiplied by the Fourier coefficients, and are added in an additional neuron. The function obtained from the additional neuron becomes the fitted curve (3).

III. SUPERVISED LEARNING METHODS OF BNN SYSTEMS

A. Supervised Learning Methods of BNN Systems Using Discrete Data

In this section, we propose a new supervised learning method of a BNN system with neurons x_1, x_2, \dots, x_N using discrete data. Fig. 4 shows the new supervised learning methods of the BNN system. As shown in Fig. 4, we use the following input and output discrete data instead of external input continuous-time functions and output continuous-time functions from neurons.

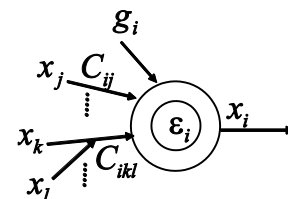


Fig. 1. Model of a neuron in a BNN system.

We define external input vectors of $2n+1$ discrete data

points as $\vec{g}_{m,i} = (g_{m,i,0}, g_{m,i,1}, \dots, g_{m,i,2n})$, and output vectors of $2n+1$ discrete data points from neurons as $\vec{x}_{m,i} = (x_{m,i,0}, x_{m,i,1}, \dots, x_{m,i,2n})$, where $i=1, 2, \dots, N$, $m=1, 2, \dots, r$, and r is the total number of external input and output vectors.

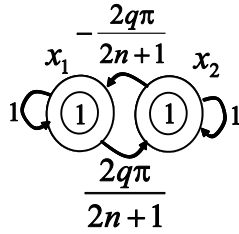


Fig. 2. The BNN system for an oscillation.

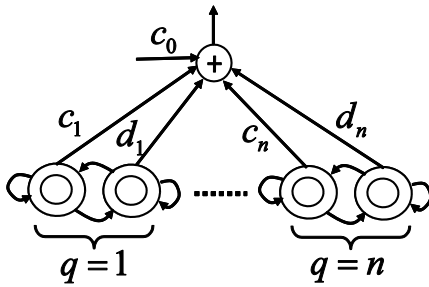


Fig. 3. The BNN system for the curve fitting composed of n BNN systems for the oscillation.

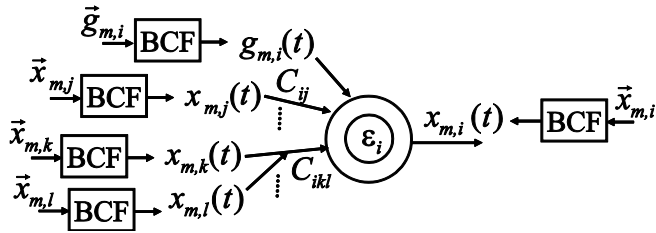


Fig. 4. The supervised learning method of a BNN system proposed in this paper when external inputs and outputs are discrete data.

The vectors of discrete data $\vec{g}_{m,i}$ and $\vec{x}_{m,i}$ are transformed into the fitted curves $g_{m,i}(t)$ and $x_{m,i}(t)$ using the BNN system for curve fitting (BCF). Then, using these fitted curves, all optimal values of C_{ij} and C_{ikl} ($i, j, k, l = 1, 2, \dots, N, k \leq l$) that satisfy (7) are determined. In the next subsection, we show how to determine all optimal values of C_{ij} and C_{ikl} using ideas taken from the method for determining all optimal values proposed in [5].

B. Method for Determining All Optimal Values

The fitted curve for $\vec{g}_{m,i}$ defined as (5) is

$$g_{m,i}(t) = \frac{1}{2n+1} \sum_{q=-n}^n G_{m,i,q} e^{\lambda q t}. \quad (8)$$

Similarly, the fitted curve for $\vec{x}_{m,i}$ is

$$x_{m,i}(t) = \frac{1}{2n+1} \sum_{q=-n}^n X_{m,i,q} e^{\lambda q t}. \quad (9)$$

Since $g_{m,i}(t)$ and $x_{m,i}(t)$ are the sums of exponential

functions, it is easier to compute derivatives of $g_{m,i}(t)$ and $x_{m,i}(t)$ compared with other functions. Moreover, since the results obtained by multiplying exponential functions are exponential functions, it becomes possible to transform the differential equation in (7) into simultaneous equations as shown below. By using $g_{m,i}(t)$ and $x_{m,i}(t)$, we compute the left-hand side and the right-hand side in (7). The left-hand side in (7) is

$$\frac{1}{2n+1} \sum_{q=-n}^n (\varepsilon_i \lambda q + 1) X_{m,i,q} e^{\lambda q t}. \quad (10)$$

The right-hand side in (7) is

$$\frac{1}{2n+1} \sum_{q=-n}^n \left(\sum_j C_{ij} X_{m,j,q} + \sum_{k \leq l} \frac{C_{ikl}}{2n+1} \left(\sum_{f(p+p')=q} X_{m,k,p} X_{m,l,p'} \right) + G_{m,i,q} \right) e^{\lambda q t}, \quad (11)$$

where we define the function $f(\omega)$ as

$$f(\omega) = \begin{cases} \omega + 2n + 1 & (\omega < -n) \\ \omega & (-n \leq \omega \leq n) \\ \omega - 2n - 1 & (n < \omega). \end{cases} \quad (12)$$

By comparing the coefficients of exponential functions in (10) with those in (11), (7) is equivalent to the following equations:

$$(\varepsilon_i \lambda q + 1) X_{m,i,q} = \sum_j C_{ij} X_{m,j,q} + \sum_{k \leq l} \frac{C_{ikl}}{2n+1} \left(\sum_{f(p+p')=q} X_{m,k,p} X_{m,l,p'} \right) + G_{m,i,q}, \quad (13)$$

where $q = -n, -n+1, \dots, n-1, n$. For $i = 1, 2, \dots, N$ and $m = 1, 2, \dots, r$, $g_{m,i}(t)$ and $x_{m,i}(t)$ satisfy (7), if and only if $g_{m,i}(t)$ and $x_{m,i}(t)$ satisfy (13). For this reason, we only have to solve the simultaneous equations (13) to determine all optimal values of C_{ij} and C_{ikl} .

Since there is the possibility that the number of simultaneous equations is larger than the number of variables, we use the method of complex linear least squares (cf. [7]) to solve the simultaneous equations. By solving the normal equation for the method of complex linear least squares (cf. [5]), we can determine all optimal values of C_{ij} and C_{ikl} .

C. Approximation of Supervised Learning Methods of BNN Systems Using Smooth Functions

As noted in the introduction, we usually cannot know the original continuous-time functions of discrete data. However, if we do know the continuous-time functions, then how is the supervised learning methods of BNN systems using continuous-time functions related to the methods proposed here? In this subsection, we show the relationship between the

supervised learning methods of BNN systems using smooth functions and our proposed methods.

When continuous-time or discrete data on a long interval are given, it is common to obtain data on a specific finite interval by applying a window function to the original data and use those data. Then the values of data on extremal points in the finite interval are zero due to the window function, and those data are periodic. Therefore we treat periodic functions and data in this section. We also use the following theorem in [5] on the approximation of a smooth periodic function and its derivatives:

Theorem 1 (cf.[5]): Let $y(t)$ be a smooth periodic function on the interval $[0,1]$. Let w be a positive integer. Then, for a sufficiently large number n , there exists a finite Fourier series on the interval $[0,2n+1]$

$$\tilde{y}(t) = \frac{1}{2n+1} \sum_{q=-n}^n Y_q e^{\lambda q t} \quad (14)$$

that satisfies

$$D^k y(t/(2n+1)) \approx D^k \tilde{y}(t) \quad (k=0,1,\dots,w). \quad (15)$$

The following is the theorem on the approximation of the supervised learning methods of BNN systems using smooth functions.

Theorem 2: Let $x_i(t)$ and $g_i(t)$ ($i=1,2,\dots,N$) be smooth periodic functions on the interval $[0,1]$, and let $\tilde{x}_i(t)$ and $\tilde{g}_i(t)$ be the finite Fourier series that satisfies (15). If $x_i(t/(2n+1))$ and $g_i(t/(2n+1))$ satisfy (7), then $\tilde{x}_i(t)$ and $\tilde{g}_i(t)$ also satisfy

$$\varepsilon_i \frac{d\tilde{x}_i}{dt} + \tilde{x}_i \equiv \sum_j C_{ij} \tilde{x}_j + \sum_{k \leq l} C_{ikl} \tilde{x}_k \tilde{x}_l + \tilde{g}_i, \quad (16)$$

when n is a sufficiently large number. Moreover, when n is increased, the right-hand side in (16) converges to the left-hand side in (16).

The proof of this theorem is given in the Appendix. By this theorem, if $x_i(t/(2n+1))$ and $g_i(t/(2n+1))$ are solutions of (7), then the n th-order approximations $\tilde{x}_i(t)$ and $\tilde{g}_i(t)$ become approximate solutions of (7). Since $\tilde{x}_i(t)$ and $\tilde{g}_i(t)$ ($i=1,2,\dots,N$) are the fitted curves composed of $2N(2n+1)$ discrete points $\tilde{x}_i(0), \dots, \tilde{x}_i(2n), \tilde{g}_i(0), \dots, \tilde{g}_i(2n)$ ($i=1,2,\dots,N$), we obtain the following corollary:

Corollary 1: Let $x_i(t)$ and $g_i(t)$ ($i=1,2,\dots,N$) be smooth periodic functions on the interval $[0,1]$. The proposed supervised learning methods, by fitted curves $\tilde{x}_i(t)$ and $\tilde{g}_i(t)$ composed of $2N(2n+1)$ discrete points, are the approximations of the supervised learning methods by $x_i(t/(2n+1))$ and $g_i(t/(2n+1))$.

Since the fitted curves approach the original smooth functions by increasing the number of discrete points,

Theorem 2 and Corollary 1 state that the optimal values of C_{ij} and C_{ikl} determined by the fitted curves also approach the optimal values of C_{ij} and C_{ikl} determined by the original smooth functions.

IV. SUPERVISED LEARNING METHODS OF BNN SYSTEMS WITH NOISE DATA

In this paper, we consider the supervised learning methods of BNN systems using discrete data with noise data, and give an example of the supervised learning methods.

We define noise vectors of $2n+1$ discrete data points as $\tilde{n}_m = (\tilde{n}_{m,0}, \dots, \tilde{n}_{m,2n})$ ($m=1, \dots, r$). By expressing the DFT component for discrete points in \tilde{n}_m as $N_{m,q}$, the fitted curve for \tilde{n}_m can be expressed as

$$n_m(t) = \frac{1}{2n+1} \sum_{q=-n}^n N_{m,q} e^{\lambda q t}. \quad (17)$$

From the definition of the curve fitting in this paper, the frequency components of \tilde{n}_m are identical to the frequency components of $n_m(t)$. Hence we understand that the curve fitting never changes the frequency components of noise. Moreover the curve fitting is linear, and changes $\tilde{g}_{m,i} + \tilde{n}_m$ and $\tilde{x}_{m,i} + \tilde{n}_m$ to $g_{m,i}(t) + n_m(t)$ and $x_{m,i}(t) + n_m(t)$ respectively.

We give one example of the proposed supervised learning methods of BNN systems using discrete data with noise data. Assume that $n=50$, $N=2$, $\varepsilon_1, \varepsilon_2=1$, and the fitted curves $g_{m,i}(t)$ and $x_{m,i}(t)$ for $\tilde{g}_{m,i}$ and $\tilde{x}_{m,i}$ satisfy

$$\frac{dx_{m,1}}{dt} + x_{m,1} = 2x_{m,1} + 4x_{m,2} + x_{m,1}x_{m,1} + 2x_{m,1}x_{m,2} + 3x_{m,2}x_{m,2} + g_{m,1} \quad (18)$$

$$\frac{dx_{m,2}}{dt} + x_{m,2} = 6x_{m,1} + 8x_{m,2} + 4x_{m,1}x_{m,1} + 5x_{m,1}x_{m,2} + 6x_{m,2}x_{m,2} + g_{m,2}, \quad (19)$$

where $m=1, \dots, r$. When noise vectors $\tilde{n}_{m,i}$ and $\tilde{n}'_{m,i}$ with random numbers generated from the normal distribution with the mean 0 and the standard deviation σ are added to $\tilde{g}_{m,i}$ and $\tilde{x}_{m,i}$, we compute optimal values of C_{ij} and C_{ikl} using $\tilde{g}_{m,i} + \tilde{n}_{m,i}$ and $\tilde{x}_{m,i} + \tilde{n}'_{m,i}$. Assume $r=1$. Then, Fig. 5 and Fig. 6 show the optimal values of C_{ij} and C_{ikl} in the case where the standard deviation σ is changed. We can confirm that the optimal values deviate from the proper values for $\sigma=0$ when σ is large. Next, we compute the optimal values in the case where r is changed, and show the results in Fig. 7 and Fig. 8. In Fig. 7 and Fig. 8, the value of σ is 0.4. As shown in Fig. 7 and Fig. 8, we can confirm that the optimal

values approach the proper values for $\sigma = 0$ by increasing the number for r .

V. CONCLUSION

We proposed supervised learning methods of BNN systems using discrete data. To transform discrete data into continuous-time functions, we introduced a curve-fitting method. The fitted curve is a smooth function that can be obtained as an output from a BNN system. Since the fitted curves can be expressed as the sum of exponential functions, it is easy to compute derivatives of the fitted curves compared with other smooth functions. Moreover, since the functions obtained by multiplying the fitted curves become the sum of exponential functions, it is possible to transform the differential equation for each neuron into simultaneous equations. By using these simultaneous equations, we proposed a method for determining the optimal values for the coefficients of all connections for each neuron in a BNN system. Moreover, we showed that the proposed supervised learning methods of BNN systems using discrete data become approximation to supervised learning methods of BNN systems using smooth functions.

When we consider supervised learning methods using smooth functions, the problem is that it is generally difficult to compute their derivatives and multiplications. If we take several discrete points to approximate these functions, we can use the fitted curves of the discrete points instead of the smooth functions themselves. Although supervised learning methods by the fitted curves become the approximations of those by the smooth functions, it is much easier to compute derivatives and multiplications of the fitted curves. Moreover, by increasing the number of discrete points, the fitted curves approach the smooth functions, so the optimal values obtained using the fitted curves approach those using the smooth functions.

As the results of experiment, we considered supervised learning methods of BNN systems using discrete data with noise data, and gave simulation results for one example of these supervised learning methods.

Currently, the BNN systems adopted in pedometers cannot be customized for each person. Therefore, we are developing a pedometer customized for each person by using the supervised learning methods of BNN systems and output data from an accelerometer embedded in a mobile phone which each person has.

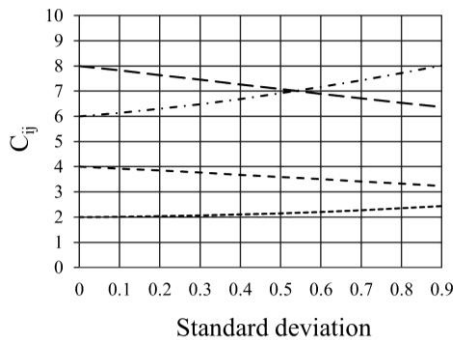


Fig. 5. The optimal values for C_{ij} in the case where the standard deviation is changed.

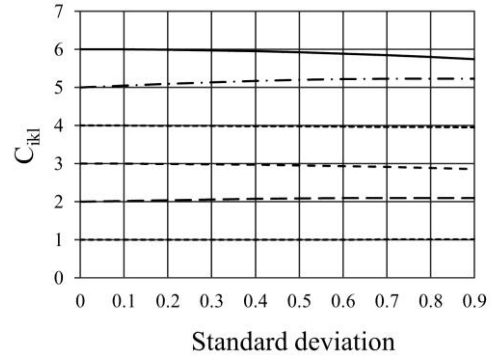


Fig. 6. The optimal values for C_{ikl} in the case where the standard deviation is changed.

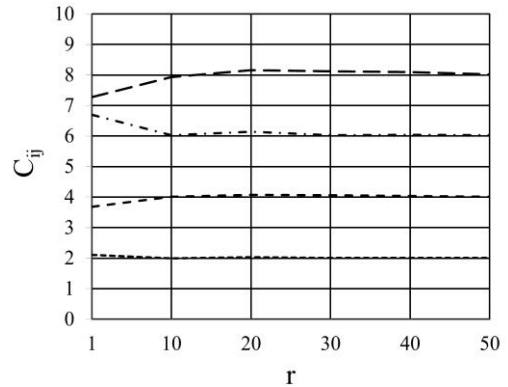


Fig. 7. The optimal values for C_{ij} in the case where r is changed.

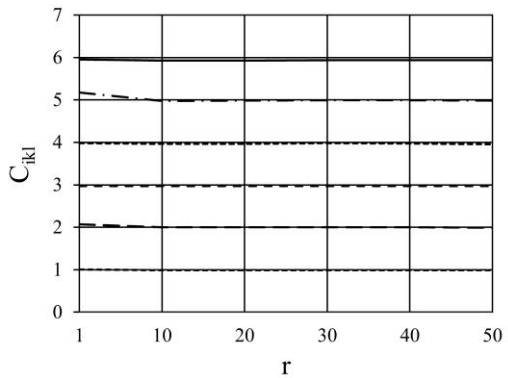


Fig. 8. The optimal values for C_{ikl} in the case where r is changed.

APPENDIX

By uniform convergence of a Fourier series (cf. [8]), the periodic functions $x_i(t/(2n+1))$ and $g_i(t/(2n+1))$ satisfy

$$x_i(t/(2n+1)) = \sum_{q=-\infty}^{\infty} \tilde{X}_{i,q} e^{\lambda_{q,t}} \tag{20}$$

$$g_i(t/(2n+1)) = \sum_{q=-\infty}^{\infty} \tilde{G}_{i,q} e^{\lambda_{q,t}} \tag{21}$$

Since the multiplication $x_i(t/(2n+1)) \cdot x_j(t/(2n+1))$ is a periodic function and hence a Fourier series, it satisfies

$$x_i(t/(2n+1)) \cdot x_j(t/(2n+1)) = \sum_{q=-\infty}^{\infty} \tilde{X}_{i,j,q} e^{\lambda_{q,t}} \tag{22}$$

and we denote the finite Fourier series of (22) as $\tilde{x}_{ij}(t)$.

From termwise differentiation of (20) (cf. [8]), we have

$$\frac{d}{dt}x_i(t/(2n+1)) = \sum_{q=-\infty}^{\infty} \lambda q \tilde{X}_{i,q} e^{\lambda q t}. \quad (23)$$

From (7), (20), (21), (22), and (23), the finite Fourier series $\tilde{x}_i(t)$, $\tilde{g}_i(t)$, $\tilde{x}_{ij}(t)$ satisfy

$$\varepsilon_i \frac{d\tilde{x}_i}{dt} + \tilde{x}_i = \sum_j C_{ij} \tilde{x}_j + \sum_{k \leq l} C_{ikl} \tilde{x}_{kl} + \tilde{g}_i. \quad (24)$$

Although $\tilde{x}_{ij}(t)$ is generally not $\tilde{x}_i(t)\tilde{x}_j(t)$, $\tilde{x}_i(t)\tilde{x}_j(t)$ converges to $\tilde{x}_{ij}(t)$ when n is increased. Therefore if n is a sufficiently large number, we obtain (16). Moreover, when n is increased, the right-hand side in (16) converges to the left-hand side in (16).

REFERENCES

- [1] S. Haykin, *Neural Networks: A Comprehensive Foundation*, NY: MacMillan College Publishing Co., 1994.
- [2] F. Nagashima, "Bilinear time delay neural network system for humanoid robot software," in *Humanoid Robots, Human-like Machines*, M. Hackel, Ed., Rijeka, Croatia: InTech, 2007, pp. 497-518.
- [3] T. Ito, Y. Senta, and F. Nagashima, "Analyzing bilinear neural networks with new curve fitting for application to human motion

- analysis," in *Proc. 2012 IEEE Int. Conf. Systems, Man and Cybernetics*, 2012, pp. 345-352.
- [4] Y. Senta, T. Ito, and F. Nagashima, "Development of pedometer for domestic dog," *ITE Tech. Rep.*, vol. 38, no. 31, pp. 63-68, Aug. 2014.
- [5] T. Ito, Y. Senta, and F. Nagashima, "Continuous-time system identification for discrete data by curve fitting," in *Proc. Int. Automatic Control Conf.*, 2015, pp. 40-47.
- [6] W. L. Briggs, and V. E. Henson, *The DFT: An Owner's Manual for the Discrete Fourier Transform*, Philadelphia: Soc. Ind. Appl. Math., 1995.
- [7] S. Cobreces, P. Rodriguez, and D. Pizarro, "Complex-space recursive least squares power system identification," in *Proc. Power Electronics Specialists Conf.*, FL, 2007, pp. 2478-2484.
- [8] K. Ito, *Encyclopedic Dictionary of Mathematics*, 2nd ed. Cambridge: MIT Press, 1993.



Toshio Ito was born in 1972. He received an M.S. degree from the Mathematical Institute at Tokyo Metropolitan University, Tokyo, Japan, in 1996 and a Ph.D. degree from the Graduate School of Mathematical Sciences at the University of Tokyo, Tokyo, Japan, in 2001.

From 2001 to 2003, he was a Researcher with the University of Tokyo. Since 2003, he has been a Researcher with Fujitsu Laboratories Ltd., Japan. From 2007 to 2008, he was a Visiting Scholar at the Center for Magnetic Recording Research, the University of California, San Diego, CA. His research interests include signal processing, artificial intelligence, and information theory.

Dr. Ito is a member of The Japanese Society for Artificial Intelligence.