A Cellular Structual Analysis for Covariance Structure

Ryo Sekiyama[†], Hisashi Aomori[†] and Mamoru Tanaka[†]

†Department of Electrical and Electronics Engineering, Sophia University,
7–1, Kioi-cho, Chiyoda-ku, Tokyo 102-8554, Japan
Phone: +81–3–3238–3878, FAX: +81–3–3238–3321
Email: r-sekiya@hoffman.cc.sophia.ac.jp

Abstract— The cellular computing concept is very powerful method for a locally connected network such like cellular neural network. Owing to its parallelization performance, this method is very suitable for a large scale problem. The covariance structure analysis is one of the well-known method for validating the given model. However, its computational cost is very high. In this paper, we propose a novel covariance structure analysis method based on cellular neural network for fast processing. Moreover, this method enables the CNN to acquire learning ability. The experiment results suggest that the proposed method is an equivalence frame work of the conventional covariance structure analysis.

1. Introduction

Machine learning methods have been used for data mining to obtain the important information and relationship from massive amount of data. Additionally, information can be analyzed visually using the passing chart. Conventionally, decision tree methods which are applicable methods for MLC++ [1], OOGDs [2] and C4.5 [3], based on so-called 'if-then' rules have been often used for data mining. However, the decision tree methods are not suitable for the information which includes continuous data.

The back propagation neural network (BPNN) is one of the effective method for learning a model based on a teacher signal. However, it is difficult to know from a theoretical viewpoint what the learning result shown about the model. Because of its problem, the covariance structure analysis becomes a one of the popular technique for finding the social and natural phenomena to realize notional relationship from statistical information. As a result, the causal relationship can be clarified from the composition concept and observation variable. Although the covariance structure analysis is the effective method for a data mining, the computational cost is high.

In this paper, we propose a CNN [4] learning algorithm based on the covariance structure analysis [5] to predict the meaningful information including continuous data. In the conventional covariance structure analysis method, since it can be thought that the endogeneous vaiable and the exogeneous variable are generated based on the respective factors, these variables are distinguished clearly. However, it can be said that these variables are just a weight coefficient between elements of the covariance structure. This enables us to introduce a cellular computing concept into the covariance structure analysis. The advantages of introducing our concept are that the covariance strcture can be regarded as a cellular connected network, its connection weight corresponds to the endogeneous or exogeneous variables, and the equilibrium solution of proposed network quaran tees the optional coefficients of the covariance structure. It is proved that the dynamics of CNN converges to the equilibrium points, if the A-template is symmetric. However, in general, coefficients of A-template for the covariance structure are asymmetric. Hence, we use the the BFGS method for solving the CNN dynamics with asymmetric A-template. The BFGS method for solving a nonlinear differential equation, are used to reduce the computational cost than Davidon-Fletcher-Powell (DFP) method. The important point is that the both methods can solve stiff systems which have large difference among eigenvalues. The resulted parameters are used as weights on edges in the signal flow graph which works for unknown input data.

2. Proposed Cellular Structural State and Measurement Equations

The model of the causal relationship is expressible in the covariance structure by two equations. The observed variable $\boldsymbol{u}$ is a known information data. The average $\boldsymbol{\mu_x}$ and the real covariance matrix $\boldsymbol{S}$ is calculated using the observed variable $\boldsymbol{u}$. The variables, $\boldsymbol{n}$ and $\boldsymbol{l}$ show the dimension of the data sets. Let $\boldsymbol{x} \in \boldsymbol{R}^n$ be the state variable vector, then the cellular structural equation is expressed by

$$\dot{\boldsymbol{x}} = -\boldsymbol{x} + \boldsymbol{A}\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{T}, \tag{1}$$

where the $\boldsymbol{f}(\boldsymbol{x})$ is nonlinear function defined by

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{cases} 1 & x \geq 1 \\ x & -1 \leq x \leq 1 \\ -1 & x \leq -1 \end{cases} \tag{2}$$

$\boldsymbol{A} \in \boldsymbol{R}^{n \times n}$ is a weight matrix which expresses connection between the state variables, and $\boldsymbol{T} \in \boldsymbol{R}^n$ is the error variable for the state variables.

The cellular measurement equation should be used to express the casual relations among the observed variables

$\boldsymbol{u} \in \boldsymbol{R}^l$ and the state variable $\boldsymbol{x}$. The measurement equation is given by

$$\boldsymbol{u} = \boldsymbol{\mu_x} + \boldsymbol{C}\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{e}, \tag{3}$$

where $\boldsymbol{C} \in \boldsymbol{R}^{l \times n}$ is a weight matrix which expresses connection between the state variables $\boldsymbol{x}$ and observed variables $\boldsymbol{u}$ for their average $\boldsymbol{\mu_x}$, and $\boldsymbol{e} \in \boldsymbol{R}^l$ is the error variables for the observed variables.

Each i-th row vector of the matrix $\boldsymbol{A}$ and $\boldsymbol{C}$ includes a weight cofficient w_{ij} on the edge from a cell from C_j to C_i. Generally, if the model is given by a signal flow graph, $\boldsymbol{A}$ and $\mathbf{C}$ become sparse matrices. Therefore, it is possible to reduce the computational cost using the sparse matrix calculation technique.

3. Optimization Method for Machine Learning in Proposed Method

In order to estimate the parameters, our proposed method uses the BFGS method and the Backward Euler method for solving a nonlinear differential equation.

3.1. Fit Function

The variables, $\boldsymbol{l}$ and $\boldsymbol{p}$ show the dimension of the data sets. Let z be the model standardized vector, then it is given by

$$z = \boldsymbol{u} - \boldsymbol{\mu_x}. \tag{4}$$

Let $\boldsymbol{S}_u \in \boldsymbol{R}^{l \times l}$ be the covariance matrix $E(zz')$ for the state variables in linear region of piece-wise linear function, then the matrix is derived theoretically from the cellular structural equation and the measurement equation as follows;

$$\begin{aligned} \boldsymbol{S}_u &= E(zz') \\ &= \boldsymbol{G}\boldsymbol{A}_0\boldsymbol{\Phi}_0\boldsymbol{A}_0'\boldsymbol{G}' \end{aligned} \tag{5}$$

where

$$\begin{aligned} \boldsymbol{G} &= (\boldsymbol{I}, \boldsymbol{O}), \\ \boldsymbol{A}_0 &= (\boldsymbol{A}, \boldsymbol{C}), \\ \boldsymbol{\Phi}_0 &= (\boldsymbol{\Delta}, \boldsymbol{\Psi}). \end{aligned} \tag{6}$$

$\boldsymbol{I}$ is the unit matrix and $\boldsymbol{O}$ is the zero matrix, $\boldsymbol{\Delta}$ and $\boldsymbol{\Psi}$ are covariance matrices of $\boldsymbol{e}$ and $\boldsymbol{T}$, and $\boldsymbol{X}'$ is a transposed matrix of $\boldsymbol{X}$.

Let $\boldsymbol{\theta} \in \boldsymbol{R}^p$ be vector of population parameters which are elements of the matrices $\boldsymbol{A}, \boldsymbol{C}, \boldsymbol{\Delta}$ and $\boldsymbol{\Psi}$, then the Generalized Least Squares (GLS) method is applied to the fit function as

$$f_{GLS}(\boldsymbol{\theta}) = \frac{1}{2}\mathrm{tr}((\boldsymbol{S} - \boldsymbol{S}_u)\boldsymbol{S}^{-1})^2, \tag{7}$$

where $\boldsymbol{S} \in \boldsymbol{R}^{l \times l}$ is the real sample covariance matrix given by

$$\boldsymbol{S} = \frac{1}{N}\boldsymbol{Z}\boldsymbol{Z}', \tag{8}$$

Table 1: Multivariate data

	Observed Item	S_1	S_2	$\ldots$	S_{50}
u_1	Color	5	3	$\ldots$	4
u_2	Style	4	4	$\ldots$	4
u_3	Power performance	4	4	$\ldots$	4
u_4	Suspension setting	3	3	$\ldots$	3

where $\boldsymbol{Z} \in \boldsymbol{R}^{l \times l}$ is the data matrix standardized by expected value from Table 1 and N is the number of samples and $\mathrm{tr}(\dot{\boldsymbol{x}})$ means a trace of the matrix $\boldsymbol{X}$.

The matrix $\boldsymbol{S}_u$ is approached to $\boldsymbol{S}$ by using optimization calculation to obtain all parameters of sparse matrices and errors. We used both the Backward Euler method and the BFGS method for solving the quasi-Newton method.

3.2. Proposed Algorithm

The function (7) should be minimized in order to determine the parameters, that is, the solution of (7) is given by

$$\mathbf{g}(\boldsymbol{\theta}) = 0, \tag{9}$$

where $\mathbf{g}(\boldsymbol{\theta}) = (\frac{f_{GLS}}{\boldsymbol{\theta}})$.

The convergency of (9) depends on its initial value. Therefore we get a solution of (9) as a solution of the Backward Euler method of the following differential equation;

$$\dot{\boldsymbol{\theta}} = \mathrm{g}(\boldsymbol{\theta}). \tag{10}$$

3.2.1. Backward Euler Method

The Backward Euler method is implicit, in which it uses the differentiation at the next time step, instead of the current one. Implicit methods are the most practical method for solving stiff systems. This method approximates the solution $f_{GLS}(\boldsymbol{\theta})$ at virtual time $t_{k+1} = t_k + h$ by solving the implicit equation:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + h\mathbf{g}(\boldsymbol{\theta}_{k+1}) \tag{11}$$

where $\mathrm{g}(\boldsymbol{\theta}_k)$ is the gradient vector evaluated at $\boldsymbol{\theta}_k$.

Since equation (11) may be nonlinear, solving it in general requires an iterative solution method. In order to suit the function (11), it solved by

$$\mathbf{F}(\boldsymbol{\theta}_k) = \boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1} - h\mathbf{g}(\boldsymbol{\theta}_k). \tag{12}$$

In this paper, quasi-Newton method is provided for solving the implicit equation.

3.2.2. Broyden-Fletcher-Goldfab-Shanno (BFGS) method

A general function $f(\boldsymbol{\theta})$ can be approximated in each iteration by a truncated Taylor series;

$$f(\boldsymbol{\theta}) \approx f(\boldsymbol{\theta}_n) + \mathrm{g}(\boldsymbol{\theta}_n)'(\boldsymbol{\theta} - \boldsymbol{\theta}_n) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_n)'\mathbf{H}(\boldsymbol{\theta}_n)(\boldsymbol{\theta} - \boldsymbol{\theta}_n), \tag{13}$$

where $\mathbf{H}(\boldsymbol{\theta}_n) \in \mathbf{R}^{p\times p}$ is the matrix of second-order partial derivatives of function with respect to $\boldsymbol{\theta}_n$. It is important to use the inverse of the Hessian matrix in our algorithm as described bellow.

However, since the Hessian leads to algorithmic and computational complexities, an approximation technique of the inverse Hessian is often used. We use the Broyden-Fletcher-Goldfab-Shanno (BFGS) method which is one of quasi-Newton methods. The update formula is as follows:

$$\mathbf{H}_{n+1} = \mathbf{H}_n + (\mathbf{1} + \frac{\mathbf{u'H_nu}}{\mathbf{z'u}})\frac{\mathbf{zz'}}{\mathbf{z'u}} - \frac{\mathbf{zu'H_n + H_nuz'}}{\mathbf{u'z}}, \quad (14)$$

where

$$\mathbf{z} = -\alpha \mathbf{H}_n \mathbf{F}(\boldsymbol{\theta}_n),$$

$$\mathbf{u} = \mathbf{F}(\boldsymbol{\theta}_{n+1}) - \mathbf{F}(\boldsymbol{\theta}_n).$$

3.2.3. Quasi-Newton Method

An implicit method requires the solution of a nonlinear equation at each time step. For one step of the Backward Euler method, we use the quasi-Newton method given by

$$\mathbf{F}(\boldsymbol{\theta}_{k+1}) = \boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k - h\mathbf{g}(\boldsymbol{\theta}_{k+1}). \quad (15)$$

In order to satisfy the equation (11), $\mathbf{F}(\boldsymbol{\theta}_{k+1})$ is solved by the Newton method.

$${}^{(n+1)}\boldsymbol{\theta}_{k+1} = {}^{(n)}\boldsymbol{\theta}_{k+1} - (\frac{\mathbf{F}({}^{(n)}\boldsymbol{\theta}_{k+1})}{{}^{(n)}\boldsymbol{\theta}_{k+1}})^{-1}\mathbf{F}({}^{(n)}\boldsymbol{\theta}_{k+1}) \quad (16)$$

$$= {}^{(n)}\boldsymbol{\theta}_{k+1} - (\mathbf{I} - h\frac{\mathbf{g}({}^{(n)}\boldsymbol{\theta}_{k+1})}{{}^{(n)}\boldsymbol{\theta}_{k+1}})^{-1}\mathbf{F}({}^{(n)}\boldsymbol{\theta}_{k+1}) \quad (17)$$

The computation of the matrix $(\frac{\mathbf{g}({}^{(n)}\boldsymbol{\theta}_{k+1})}{{}^{(n)}\boldsymbol{\theta}_{k+1}})$ is very difficult task. Here, the approximation technique BFGS is used. We replace $(\mathbf{I} - h\frac{\mathbf{g}({}^{(n)}\boldsymbol{\theta}_{k+1})}{{}^{(n)}\boldsymbol{\theta}_{k+1}})$ by the approximation (Eq. (14)). The flowchart of proposed method is given in Fig. 1.

4. Simulation Results

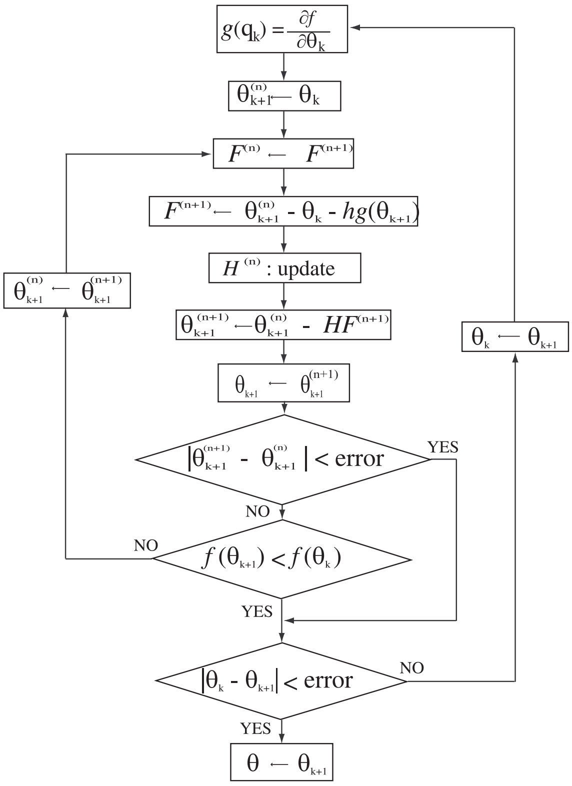
The model of "Purchase of a Car" is used as an example of analysis of the cellular covariance structure. The observed data sets are shown in Table data.

This data sets were collected from a survey of 50 people. The four observed variables u_1, u_2, u_3, u_4 are defined as observed variables in Table 1.

The state variable of x_1 means a design, the state variable of x_2 means a performance, and the state variable of x_3 means a value of a car. The parameters of ζ_1, ζ_2 and ζ_3 are the error variables.

The SFG corresponding to the basic equations is given in Fig. 2.

A design and a performance are determined by a value of a car. A user can set the edges in advance. Some parameters of the matrices are set to 0 before learning. It is useful to set some parameters previously, if possble.



$g(q_k) = \frac{\partial f}{\partial \theta_k}$
$\theta^{(n)}_{k+1} \leftarrow \theta_k$
$F^{(n)} \leftarrow F^{(n+1)}$
$F^{(n+1)} \leftarrow \theta^{(n)}_{k+1} - \theta_k - hg(\theta_{k+1})$
$H^{(n)}$: update
$\theta^{(n)}_{k+1} \leftarrow \theta^{(n+1)}_{k+1}$
$\theta^{(n+1)}_{k+1} \leftarrow \theta^{(n)}_{k+1} - HF^{(n+1)}$
$\theta_k \leftarrow \theta_{k+1}$
$\theta_{k+1} \leftarrow \theta^{(n+1)}_{k+1}$
$|\theta^{(n+1)}_{k+1} - \theta^{(n)}_{k+1}| <$ error
YES
NO
NO
$f(\theta_{k+1}) < f(\theta_k)$
YES
$|\theta_k - \theta_{k+1}| <$ error
NO
YES
$\theta \leftarrow \theta_{k+1}$

Figure 1: Flowchart of proposed method

It is very important that the coefficient matrices are sparse and the SFG is a cellular network. The weights on the edges incident to a cell C_j are corresponding to the template like that of a cellular neural network.

The cellular structural state equation of the model for "Purchase of a Car" is given by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & \beta_{12} & \beta_{13} \\ 0 & 0 & \beta_{23} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{pmatrix} \quad (18)$$

The cellular measurement equation can be also defined by the user as follows

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} + \begin{pmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & 0 & 0 \\ 0 & \kappa_{32} & 0 \\ 0 & \kappa_{42} & 0 \end{pmatrix} \begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix} \quad (19)$$

Fig. 3 and Fig. 4 show the learning curves for the model of "Purchase of a Car". The number of steps for Backward Euler method is shown on the horizontal axis. The

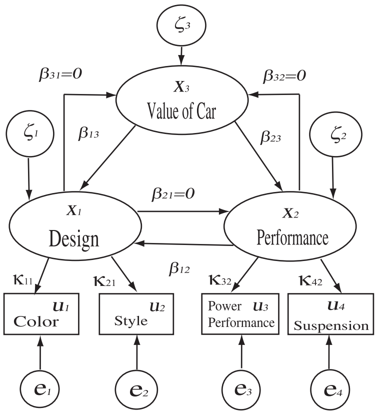


Figure 2: SFG of ”Purchase of a Car” model

fit function, derived such that the covariance matrix C_u is approached to S is shown on the vertical axis.

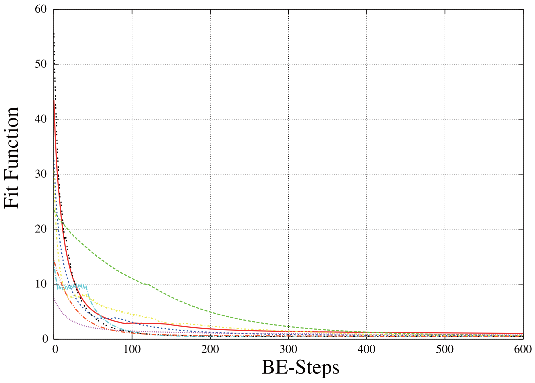


Figure 3: The simulation result by the DFP method.

Fig. 5 shows the comparison result of estimate parameters between proposed and original method [5]. The parameters are shown on the horizontal axis. The value of each parameter is shown on the vertical axis.

5. Conclusion

In this paper, a cellular structual analysis method for covariance structure has been proposed. The proposed CNN structure corresponds to the given model and its dynamics is utilized for estimating the parameters that describe the given model. The experimental results show that our method is an equivalent method of the conventional covariacne structure analysis with fast computational performance.

Acknowledgments

This research is supported by the fund of Open Research Center Project from MEXT of Japanese Government (2007-2011).

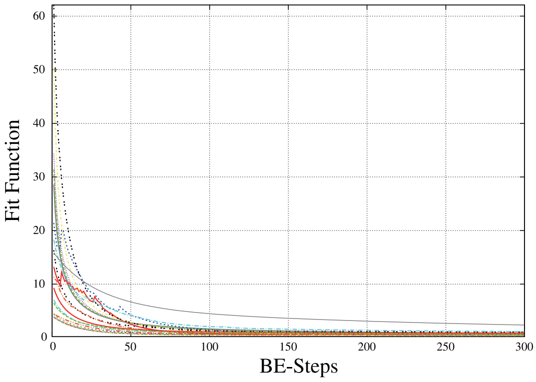


Figure 4: The simulation result by the BFGS method.

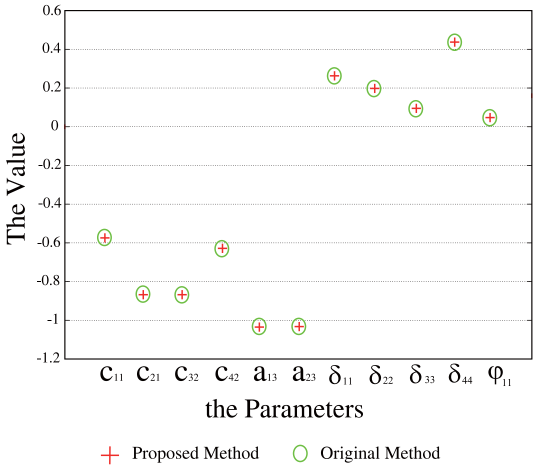


Figure 5: Estimate Parameters.

References

[1] R. Kohavi, D. Sommerfield, and J. Dougherty, ”Data Mining using MLC++ : A machine learning library in C++”, International Journal of Artificial Intelligence Tools, Vol. 6, No. 4, pp. 537-566. 1997.

[2] R. Kohavi, ”Wrappers for Performance Enhancement and Oblivious Decision Graphs”, Ph.D. dissertation Stanford, CA: Comp. Sci. Department of Computer Science, Stanford University 1995.

[3] Quinlan, J. Ross, ”C4.5: Programs for machine learning”, Morgan Kaufmann Publishers, 1993.

[4] L. O. Chua and L. Yang, “Cellular neural networks: theory,” *IEEE Trans. Circuits Syst.,* vol. 35, no. 10, pp. 1257–1272, Oct. 1988.

[5] Y. Zennyoji, N. Ohashi, M. Yamauchi, and M. Tanaka, ”Cellular Analysis of Covariance Structure for Data Mining by Backward Euler Method,”In Proc. 2005 Intl. Symp. Nonlinear Theory and its Appl. (NOLTA 2005), Bruges, Belgium, Oct. 2005.