

ニューラルネットワークによる線形構造系の同定

佐藤 忠信・佐藤 誠

IDENTIFICATION OF LINEAR STRUCTURAL SYSTEM BY NEURAL NETWORK

By *Tadanobu SATO* and *Makoto SATO*

Synopsis

This paper explores the potential of using neural network approaches to identify the dynamic characteristics of structural system. The relevant neural network characteristics of learning algorithm is discussed in the context of system identification. Because of self-learning nature of neural network the identified dynamic characteristics are strongly affected by the level of noise contained in the teaching signals. Using the Karman filtering technique, a method to identify the dynamic characteristics of structural system proof against contaminating noise in teaching signals is developed.

1. はじめに

近年、土木構造物の長大化、高層化が著しく進み、それに伴って構造物の地震や風による揺れを抑えることが必要不可欠になってきている。最近では地震に耐えるというような受動的な思想から脱却し、より能動的に揺れを抑えようとする考え方が台頭しており、機械工学や航空工学の分野で発達してきた最適制御則を用いて、構造物の揺れを抑えるアクティブ制振の技術が実用段階を迎えている。これは対象となる構造物に外部から何らかの方法でその振動を抑止しようとする力を加え、揺れを低減しようとするもので、その制御系を設計するためには対象となる構造物の動特性を正確に捕えることが、制御効率の向上を図るためばかりでなく、制御対象の安全性を保証するためにも重要な課題となっている。

従来より構造物が地震などの外乱に対して示す力学的挙動を推定するための物理モデルを決定する試みは多くなされてきた。土木構造物の地震や風に対する力学的特性は、部材の質量や剛性によってのみ決定されているのではなく、構造形式などにも依存しているため、これまで構造物の固有周期や減衰定数といった値は、使用されている部材についての情報と、すでに建設された構造物によって経験的に得たデータに基づいて推定されることが多かった。しかしながらこうして推定される構造パラメータには誤差が含まれ、構造物の挙動が正確に把握できないことがあった。また供用後の部材の劣化にともなう強度の低下や、基礎の圧密沈下による不等沈下の発生などによって、対象となる構造物の動特性が変動するので、対象となる構造物について一定期間ごとにその動的特性を同定することも必要である。

対象とする構造物の挙動の観測データから、その動特性を表わす物理的なモデルを決定したり、モデルの特性を規定しているパラメータを決定することをシステム同定という。観測データをもとに構造物の同定を行うときの最も重要な課題は観測値に含まれるノイズをいかに削除し、誤差の少ない値を決定するかという点である。

動特性の推定には、Winner による定常信号理論を用いた周波数領域での解析がある。この方法は、観測波形を周波数領域に変換し、そのスペクトルの様子から動特性を推定するものである。しかし減衰定数や高

次モードの固有振動数を推定するのは非常に困難であった。そこで Kalman¹⁾²⁾ は観測波形を、時間領域で扱うことのできる状態空間表示をもとにした線形システム理論を提唱した。これは状態変数法あるいは実時間法と呼ばれるもので、ばらつきの多い観測データをもとに同定を行っても、ノイズ成分を削除し、精度よくパラメータの推定が行うことができる。このノイズ成分を削除する性質から、Kalman の提唱したこの手法は、カルマンフィルタと呼ばれ、後にこれを非線形システムにも適用した拡張カルマンフィルタのアルゴリズムも開発され、現在幅広い分野で活用されている³⁾⁴⁾⁵⁾⁶⁾。土木工学でも星谷ら⁷⁾⁸⁾ はカルマンフィルタを用いて各種の振動系のパラメータ推定を行い、また、土岐ら⁹⁾ は非線形復元力特性の同定を行っている。カルマンフィルタに代表される最小2乗法や最尤法といった枠組みで線形システムを扱う限りでは、システム同定は1980年代なかばには理論的な完成を見たと考えられる。しかしそれをいかに実システムに適用するかという実際的な問題が残されており、今もなお活発な研究が行なわれている¹⁰⁾。

システム同定に関する研究成果は Beck¹¹⁾ や Natke¹²⁾ によって要約されている。しかし、これまでに開発されてきた手法を用いるだけでは、アルゴリズムの安定性が不十分であったり、計算時間の問題などで大規模なシステムを効率よく同定することが困難になることが多かった。こうした問題点を解決するために、高いレベルの並列処理が可能な同定アルゴリズムの開発が求められるようになってきている¹³⁾。

一方、情報工学の分野では人間の脳で行われる複雑かつ膨大な量の情報処理のメカニズムにヒントを得て、神経回路網状のネットワークメカニズムによる情報処理への可能性を探る試みが行われてきた¹⁴⁾。このネットワークをニューラルネットワークと呼ぶ。この研究は1943年の McCulloch-Pitts¹⁵⁾ のニューロンモデルの研究に始まり、1957年、Rosenblatt¹⁶⁾ による学習するパターン識別機械、パーセプトロンの研究に刺激され活発化した。人間の脳による脳の理解、人体の行っている情報処理、脳のような複雑な非線形システム、ひとりでに学習し自己組織していくシステム、高い並列性を持った情報処理、というような様々な興味の対象として、情報工学のみならず、医学、心理学、などの幅広い研究分野においてニューラルネットワークの研究が行なわれた。その後しばらくニューラルネットワークの研究は停滞するが、1986年の Rummelhart¹⁷⁾ による階層型ネットワークに対する誤差逆伝播法 (Back Propagation; BP) のアルゴリズムの提唱により、あらためて活発に応用研究が行われるようになり、この理論は実用段階に達している。機械工学や電気工学の分野などではニューラルネットワークの持つ自己学習機能、非線形写像能力を生かして、ロボットの運動をコントロールしたり、音声や文字を解読するシステムの開発などが進められている。しかし研究が進むにつれ、誤差逆伝播法の持つ問題点も明確になってきた¹⁸⁾。その問題点としては

- (1) 学習に用いる教師信号にノイズが混入していると学習過程が不安定になる
- (2) 学習の初期条件によっては解が局所解に陥る危険性がある
- (3) 2乗出力誤差評価が単に入出力間の写像の再現を目的としているため、ネットワークの内部がブラックボックス的になりその解釈が困難となる
- (4) 最大傾斜法を用いているため最適点に近づくに従い評価関数の勾配が小さくなり、一回の修正量が小さくなるので学習に時間がかかる

などがあり、これまでもこのような問題点を解決するためにさまざまな研究がなされてきた。ニューラルネットワークを同定問題に適用するためには、序論で触れたようにノイズが混入された観測データであっても、安定した学習が行なえるような学習アルゴリズムが必要である。

このようなことを踏まえ、本研究では、ニューラルネットワーク理論の持つ高度な並列処理能力、非線形写像能力に着目し、ニューラルネットワーク理論をシステム同定に適用することで、システム同定が必要としてきた高いレベルの並列処理を実現する。そしてシステム同定の最重要課題であるノイズに強いアルゴリズムを作るため、学習法に従来の方法ではなく拡張カルマンフィルターを用いたノイズに強い学習法を提案し、より効率的な同定アルゴリズムの構築を図る。

2. ニューラルネットワーク

ニューラルネットワークは神経回路網での複雑な情報の処理を簡単なネットワークモデルに置換し、そのモデル上で対象とする情報処理系を再現するものである。ニューラルネットワークには数々のモデルが提案されているが、最も簡単に幅広く用いられているものに Perceptron 型と呼ばれる階層型ネットワークがあり、Fig. 1 に示されるような構造を有している。Fig. 1 は 3 層よりなるネットワークの例であり、Perceptron モデルでは入力層→中間層→出力層の順番に一方方向のみの結合を持ち、信号は入力層から出力層へ一方向に伝わる。

基本的な信号の伝播の様子は以下のようなものである。ある層のノードへの入力が前層の各ノードからの信号の重み付きの和として表現され、それが非線形入出力関数を通して出力され、次の層へ入力されるというシステムである。

ここでニューラルネットワーク内での信号の処理過程を Fig. 1 のネットワークをもとに示す。まず本解析で用いる記号を以下のように定義する。各層でのそれぞれのノードへの入力および出力を

$u = \{u_1, u_2, \dots, u_n\}^T$	$; n \times 1$	入力層への入力ベクトル
$\bar{v} = \{\bar{v}_1, \bar{v}_2, \dots, \bar{v}_p\}^T$	$; p \times 1$	第 1 層への入力ベクトル
$v = \{v_1, v_2, \dots, v_p\}^T$	$; p \times 1$	第 1 層からの出力ベクトル
$\bar{z} = \{\bar{z}_1, \bar{z}_2, \dots, \bar{z}_q\}^T$	$; q \times 1$	第 2 層への入力ベクトル
$z = \{z_1, z_2, \dots, z_q\}^T$	$; q \times 1$	第 2 層からの出力ベクトル
$\bar{y} = \{\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m\}^T$	$; m \times 1$	第 3 層への入力ベクトル
$y = \{y_1, y_2, \dots, y_m\}^T$	$; m \times 1$	第 3 層からの出力ベクトル

とし、各層の間を結んでいるリンクの重みを次の 3 個のマトリクスで表現する。

$w^1 = \{w_{ij}^1\}$	$; p \times n$	入力層 j 番目のノードと第 1 層 i 番目のノードを結ぶ重み
$w^2 = \{w_{ij}^2\}$	$; q \times p$	第 1 層の j 番目のノードと第 2 層の i 番目のノードを結ぶ重み
$w^3 = \{w_{ij}^3\}$	$; m \times q$	第 2 層の j 番目のノードと第 3 層の i 番目のノードを結ぶ重み

信号の伝播の様子は以下のようなものである。まず入力層に $\{u\}$ が入力されるとそのそれぞれの要素の値に各リンクの重み係数 w_{ij}^1 を乗じ、それらの値の総和が第 1 層内の各ノードへの入力となる。即ち

$$\bar{v}_i = \sum_{j=1}^n (w_{ij}^1 \cdot u_j) \dots \dots \dots (1)$$

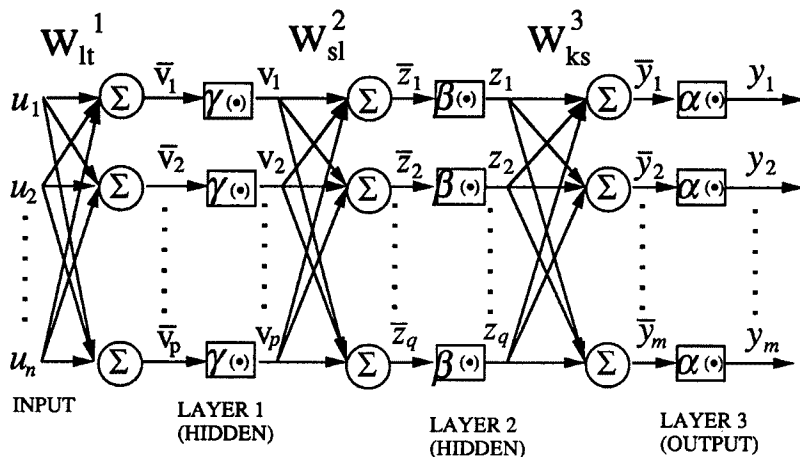


Fig. 1 Perceptron neural network model.

であり、マトリクス表現をすると

$$\{\bar{v}\} = [w^1] \{u\} \dots \dots \dots (2)$$

となる。次に第1層のそれぞれのノードに設定された非線形入出力関数を経て第1層からの出力値となる。

$$v_i = \gamma(\bar{v}_i) \dots \dots \dots (3)$$

以下第2層、第3層でもおなじような処理が施され、次式のように変換される。

第2層において

$$\{\bar{z}\} = [w^2] \{v\} \quad z_i = \beta(\bar{z}_i) \dots \dots \dots (4)$$

第3層において

$$\{\bar{y}\} = [w^3] \{z\} \quad y_i = \alpha(\bar{y}_i) \dots \dots \dots (5)$$

このようにしてネットワークに入力が与えられたときの出力 y_k が計算できる。この y_k は重み係数 w_{ij}^1 , w_{ij}^2 及び w_{ij}^3 の非線形関数となる。ニューラルネットワークでは層間の結合の重み係数 w_{ij}^1 , w_{ij}^2 と w_{ij}^3 がネットワークの入力—出力関係を目標の入力—出力のペアに合致するように修正されていくという特徴がある。目標となる入力—出力のペアを教師信号と呼び、重みを修正していく過程は学習と称される。以下では教師信号として地震時に構造物が示す応答の時刻歴を用いて、地震時に構造物が示す時刻での応答値を入力とし、微小時間 dt 後の応答値を出力とするネットワークを構築する。

3. 誤差逆伝播法による学習

甘利によって発見され、Rumelhartら¹⁷⁾によって再発見された誤差逆伝播法による学習は階層型ネットワークで、直接出力層につながっていないリンクの重みも学習できるため、幅広く活用されてきた。誤差逆伝播のアルゴリズムでは、式(6)で与えられる教師信号と出力値との2乗誤差によって表現された評価関数を最小にするように、最大傾斜法を用いて式(7)に基づいてネットワークの重み係数が更新される。

$$E = \frac{1}{2} \sum_{k=1}^m e_k^2 \quad e_k = y_k - y_{kd} \dots \dots \dots (6)$$

$$w^{(new)} = w^{(old)} - \varepsilon \cdot \frac{\partial E}{\partial w^{(old)}} \dots \dots \dots (7)$$

ここに y_k はネットワークからの出力値、 y_{kd} は観測値、 ε は学習率と呼ばれる小さな正の値で、評価関数の勾配に乗せられるパラメータであり、一回の学習で修正される重みを決定し、学習の進行を左右する。式(7)の右辺第2項にある偏微分係数は次の様に表される。

$$\nabla_3 E = \left\{ \frac{\partial E}{\partial w_{ij}^3} \right\} = \begin{pmatrix} \delta_1^3 z_1 & \dots & \delta_1^3 z_q \\ \vdots & \ddots & \vdots \\ \delta_m^3 z_1 & \dots & \delta_m^3 z_q \end{pmatrix} \quad \delta_i^3 = e_i \alpha'(\bar{y}_i) \dots \dots \dots (8)$$

$$\nabla_2 E = \left\{ \frac{\partial E}{\partial w_{ij}^2} \right\} = \begin{pmatrix} \delta_1^2 v_1 & \dots & \delta_1^2 v_p \\ \vdots & \ddots & \vdots \\ \delta_q^2 v_1 & \dots & \delta_q^2 v_p \end{pmatrix} \quad \delta_i^2 = \beta(\bar{z}_i) \cdot \sum_{k=1}^m w_{ki}^3 \delta_k^3 \dots \dots \dots (9)$$

$$\nabla_1 E = \left\{ \frac{\partial E}{\partial w_{ij}^1} \right\} = \begin{pmatrix} \delta_1^1 u_1 & \dots & \delta_1^1 u_n \\ \vdots & \ddots & \vdots \\ \delta_p^1 u_1 & \dots & \delta_p^1 u_n \end{pmatrix} \quad \delta_i^1 = \gamma'(\bar{v}_i) \cdot \sum_{k=1}^q w_{ki}^2 \delta_k^2 \dots \dots \dots (10)$$

また誤差逆伝播という名の由来は、式 (8) から (10) において重み係数の修正量を計算するときに現われる δ がそれより一つ先の層の重み係数を修正するため求めた δ の重み付の和として表現され、その様子が最終層の出力と教師信号の差があたかもネットワークの中を、出力層から入力層に向かって逆に伝播しているかのごとく重みが修正されていることによるものである。

4. カルマンフィルタを用いた学習

前書きのところで述べた問題点を解決すべく、拡張カルマンフィルタを用いて重み係数の学習がおこなえるアルゴリズムを誘導する。カルマンフィルタはベイズの推定に基づく統計的推定法である。平均値、標本分散などの値やこれらと共に同一条件下のデータを母集団といい、母集団の特性を示す真のパラメータを母数と言うが、統計的推定法は母集団のごく限られたデータを用いてその母集団の母数を推論しようとするものである。またベイズの推定は母数を確率変数すなわち推定誤差を確率量として推定する方法である。Fig. 1 のネットワークでは出力 y_k は重み係数 $w_{ij}^1, w_{ij}^2, w_{ij}^3$ の非線形関数となっているので、まずその代表値 $\hat{w}_{ij}^1, \hat{w}_{ij}^2, \hat{w}_{ij}^3$ のまわりで線形化する。即ち出力を $(w_{ks}^3 - \hat{w}_{ks}^3)$ と $(w_{si}^2 - \hat{w}_{si}^2)$ ならびに $(w_{ij}^1 - \hat{w}_{ij}^1)$ の関数に展開しなければならない。まず、第3層における重み係数に関する線形化をテラー展開の1次のオーダーまでを用いることによって、次式のように表わす。

$$y_k = \alpha(\bar{y}_k) + \frac{\partial \alpha}{\partial \bar{y}_k} \frac{\partial \bar{y}_k}{\partial w_{ks}^3} (w_{ks}^3 - \hat{w}_{ks}^3) = \alpha(\bar{y}_k) + \alpha'(\bar{y}_k) \sum_{s=1}^q z_s \cdot (w_{ks}^3 - \hat{w}_{ks}^3) \dots (11)$$

この式をマトリクスの形式で表現すれば

$$\{y_k\} = \{\alpha(\bar{y}_k)\} + \begin{bmatrix} \alpha'(\bar{y}_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha'(\bar{y}_m) \end{bmatrix} [w_{ks}^3] \{z_s\} - \begin{bmatrix} \alpha'(\bar{y}_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha'(\bar{y}_m) \end{bmatrix} [\hat{w}_{ks}^3] \{z_s\} \dots (12)$$

と表現され、さらに右辺第2項と3項に現われる重み係数マトリクスを次式のような重み係数ベクトルに置き換える。

$$\{w^3\} = \{w_{ij}^3\}^T \{w_{ij}^3\}^T \dots \{w_{ij}^3\}^T \quad (j=1, 2, \dots, q) \dots (13)$$

この場合、式 (12) の右辺第2項は次式のように表される。

$$\begin{bmatrix} \alpha'(\bar{y}_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha'(\bar{y}_m) \end{bmatrix} [\hat{w}_{ks}^3] \{z_s\} = [A_3] \{w^3\} \dots (14)$$

ここに $[A_3]$ は kk 成分以外はゼロのマトリクスで、その kk 成分の表現は、

$$A_3^{kk} = \{\alpha'(\bar{y}_k) z_1 \quad \alpha'(\bar{y}_k) z_2 \quad \dots \quad \alpha'(\bar{y}_k) z_q\} \dots (15)$$

で与えられる。式 (12) の右辺第3項についても、 $\{w_{ks}^3\}$ を $\{\hat{w}_{ks}^3\}$ に置き換えれば式 (14) と同じような表現ができる。

次に、第2層における重み係数に対する線形化を行う。第2層の重み係数に関する第1近似は式 (11) の右辺第1項のみをテラー展開し、その1次のオーダーまでを用いれば求められるので

$$\alpha(\bar{y}_k) = \alpha\left(\sum_{s=1}^q \hat{w}_{ks}^3 z_s\right) = \alpha\left(\sum_{s=1}^q w_{ks}^2 \hat{z}_s\right) + \frac{\partial \alpha}{\partial \bar{y}_k} \frac{\partial \bar{y}_k}{\partial z_s} \frac{\partial z_s}{\partial \hat{z}_s} \frac{\partial \hat{z}_s}{\partial w_{si}^2} (w_{si}^2 - \hat{w}_{si}^2) \dots (16)$$

となり、従って

$$y_k = \alpha\left(\sum_{s=1}^q \hat{w}_{ks}^3 \hat{z}_s\right) + \alpha'\left(\sum_{s=1}^q \hat{w}_{ks}^3 \hat{z}_s\right) \sum_{s=1}^q \hat{z}_s \cdot (w_{ks}^3 - \hat{w}_{ks}^3) + \alpha'(\bar{y}_k) \sum_{s=1}^q \hat{w}_{ks}^3 \beta'(\bar{z}_s) \sum_{l=1}^p v_l \cdot (w_{si}^2 - \hat{w}_{si}^2) \dots (17)$$

となる。これもマトリクスの形式で表現すれば

$$\{y_k\} = \{\alpha(\hat{y}_k)\} + [\alpha'(\hat{y}_k)]\{[w_{ks}^2] - [w_{ks}^1]\}\{z_s\} + [\alpha'(\hat{y}_k) \hat{w}_{ks}^3 \beta'(\bar{z}_s)]\{[w_{si}^2] - [\hat{w}_{si}^2]\}\{v_i\} \quad (18)$$

と表わされ、式 (13) と同じく次のような重み係数ベクトルを定義する。

$$\{w^2\} = \left\{ \{w_{1j}^2\}^T \{w_{2j}^2\}^T \cdots \{w_{pj}^2\}^T \right\}^T \quad (j=1, 2, \dots, p) \quad \dots\dots\dots (19)$$

この場合、式 (18) の右辺第 3 項は次式のようになる。

$$[\alpha'(\hat{y}_k) \hat{w}_{ks}^3 \beta'(\bar{z}_s)]\{[w_{si}^2] - [\hat{w}_{si}^2]\}\{v_i\} = [A_2]\{\{w^2\} - \{\hat{w}^2\}\} \quad \dots\dots\dots (20)$$

ここにマトリクス $[A_2]$ の ks 成分は次のようになる。

$$A_2^{ks} = \{F_{ks}v_1 \quad F_{ks}v_2 \quad \cdots \quad F_{ks}v_p\} \quad \dots\dots\dots (21)$$

ただし

$$F_{ks} = \alpha'(\hat{y}_k) \hat{w}_{ks}^3 \beta'(\bar{z}_s)$$

次に、第 1 層における重み係数に対する線形化を行う。第 1 層の重み係数に関する第 1 近似は式 (18) の右辺第 1 項のみをテーラー展開し、その 1 次のオーダーまでとれば求められるので

$$\begin{aligned} \alpha(\hat{y}_k) &= \alpha\left(\sum_{s=1}^q w_{ks}^2 \beta\left(\sum_{i=1}^k w_{si}^2 \hat{v}_i\right)\right) + \frac{\partial \alpha}{\partial \hat{y}_k} \frac{\partial \hat{y}_k}{\partial z_s} \frac{\partial z_s}{\partial \bar{z}_s} \frac{\partial \bar{z}_s}{\partial v_i} \frac{\partial v_i}{\partial \bar{v}_i} \frac{\partial \bar{v}_i}{\partial w_{si}^2} (w_{si}^2 - \hat{w}_{si}^2) \\ &= \alpha(\hat{y}_k) + \alpha'(\hat{y}_k) \sum_{s=1}^q w_{ks}^2 \beta'(\bar{z}_s) \sum_{i=1}^k \hat{w}_{si}^2 \gamma'(\hat{v}_i) \sum_{i=1}^n u_i (w_{si}^2 - \hat{w}_{si}^2) \quad \dots\dots\dots (22) \end{aligned}$$

となり、この右辺第 2 項をマトリクス形式で書けば

$$[\alpha'(\hat{y}_k) \hat{w}_{ks}^3 \beta'(\bar{z}_s)] [\hat{w}_{si}^2 \gamma'(\hat{v}_i)]\{[w_{si}^2] - [\hat{w}_{si}^2]\}\{u_i\} \quad \dots\dots\dots (23)$$

と表されるので、

$$A_1 = \begin{bmatrix} A_1^{11} & \cdots & A_1^{1l} & \cdots & A_1^{1p} \\ \vdots & & \vdots & & \vdots \\ A_1^{k1} & \cdots & A_1^{kl} & \cdots & A_1^{kp} \\ \vdots & & \vdots & & \vdots \\ A_1^{m1} & \cdots & A_1^{ml} & \cdots & A_1^{mp} \end{bmatrix} \quad \dots\dots\dots (24)$$

ただし

$$\begin{aligned} A_1^{kl} &= \{G_{kl}u_1 \quad G_{kl}u_2 \quad \cdots \quad G_{kl}u_n\} \\ G_{kl} &= \alpha'(\hat{y}_k) \sum_{s=1}^q \hat{w}_{ks}^3 \beta'(\bar{z}_s) \sum_{i=1}^k \hat{w}_{si}^2 \gamma'(\hat{v}_i) \end{aligned}$$

を定義した上で、式 (13) と同じように式 (25) のような重み係数ベクトルを定義し全体を重ね合わせれば

$$\{w^1\} = \left\{ \{w_{1j}^1\}^T \{w_{2j}^1\}^T \cdots \{w_{pj}^1\}^T \right\}^T \quad (j=1, 2, \dots, n) \quad \dots\dots\dots (25)$$

$$\{y_k\} = \left\{ \alpha(\hat{y}_k) \right\} + [A_3]\{w^3\} - [A_3]\{\hat{w}^3\} + [A_2]\{w^2\} - [A_2]\{\hat{w}^2\} + [A_1]\{w^1\} - [A_1]\{\hat{w}^1\} \quad \cdots (26)$$

と表現できる。ここに $\left\{ \alpha(\hat{y}_k) \right\}$ は前ステップでの重み係数を用いて出力された値である。

式 (26) の重み係数ベクトル $\{w\}$ をカルマンフィルターにより決定するために、次のベクトルとマトリクスを定義する。

$$\{w_t\} = \left\{ \{w^3\} \{w^2\} \{w^1\} \right\} \dots\dots\dots (27)$$

$$[H_t] = \left[[A_3] [A_2] [A_1] \right] \dots\dots\dots (28)$$

$$\{h_t\} = \left\{ \alpha(\hat{y}_k) \right\} \dots\dots\dots (29)$$

観測ベクトルを y_t , 観測ノイズを v_t とおけば式 (26) より観測方程式は

$$y_t = h_t + H_t w_t - H_t \hat{w}_t + v_t \dots\dots\dots (30)$$

システム方程式は

$$w_{t+1} = I w_t \dots\dots\dots (31)$$

と表現され、拡張カルマンフィルターのアルゴリズムが次の1) から6) で与えられる。ここに $t/t-1$ の添字のついた変量は離散時間の第 $t-1$ ステップ時における第 t ステップ時に対する推定量を表し、 P は重み係数の推定誤差の共分散、 R は測定誤差の共分散、 Q は系に加わる外乱の共分散である。

1) 重み係数ベクトルの推定値の計算

$$\hat{w}_{t+1/t} = I \hat{w}_{t/t} \dots\dots\dots (32)$$

2) 推定誤差共分散マトリクス of 計算

$$P_{t+1/t} = P_{t/t} + Q_t \dots\dots\dots (33)$$

3) カルマンゲインの計算

$$K_t = P_{t/t-1} H_t^T [H_t P_{t/t-1} H_t^T + R_t]^{-1} \dots\dots\dots (34)$$

4) 重み係数ベクトルの更新

$$\hat{w}_{t/t} = \hat{w}_{t/t-1} + K_t [y_t - h_t(w_{t/t-1})] \dots\dots\dots (35)$$

5) 推定誤差共分散マトリクス of 更新

$$P_{t/t} = P_{t/t-1} + H_t^T [H_t P_{t/t-1} H_t^T + R_t]^{-1} H_t P_{t/t-1} \dots\dots\dots (36)$$

6) 第1ステップへ

5. 解析結果

5.1 解析の条件

まず、Fig. 2 に示す 2 自由度線形モデルの地震波に対する応答計算結果を観測されるデータ、即ち教師信号としてネットワーク中の重み係数の学習を行う。モデルの諸元値は Table 1 に示した値とする。使用するネットワークモデルは Fig. 3 に示す 2 層構造とし、設定する入出力関数は全て線形関数とする。そして時刻 k での変位応答・速度応答および地動加速度を入力し、時刻 $k+1$ の変位、速度の応答を出力するネットワークを構築する。この場合、重み係数は 4×5 のマトリクスとなる。また時間間隔の刻みは 0.01 秒とした。

5.2 推定誤差共分散行列の初期値の同定結果に及ぼす影響

カルマンフィルタを用いて重み係数の学習を行う場合には、重み係数の推定誤差共分散行列 P 及び観測

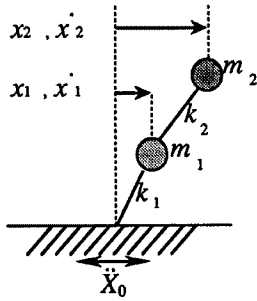


Fig. 2 Structure model.

Table 1. Model parameters.

Mass (kgs ² /cm)	2F	0.12553
	1F	0.12553
Stiffness (kgf/cm)	1F-2F	24.5
	1F-Base	24.5
Damping (%)	h ₁	0.863
	h ₂	2.260

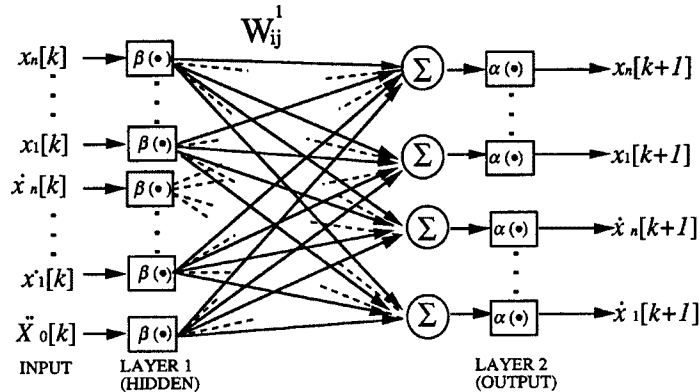


Fig. 3 2-Layered neural network.

誤差の共分散 R の初期値を与える必要がある。本解析では重み係数が合計 20 個であるので P は 20×20 の次元のマトリクス、観測データとしては 2 つの質点の変位及び速度がそれぞれ与えられるので、 R は 4×4 のマトリクスとなる。これらの共分散は、同定を行う際に用いる観測データを測定した計測器の精度や、先見的情報として得られている構造パラメータの初期値の正確さを十分に考慮した上で設定する必要があるが、本解析では P の初期値として対角マトリクスを仮定し、モデルの状態方程式を解くことによって得られる重み係数の厳密解の 2 乗に比例定数 p を乗じて与える。観測データに含まれるノイズ成分の共分散行列 R の初期値についても対角マトリクスを仮定し、モデルの 1 次モード形の 2 乗に比例定数 r を乗じた形で与えるものとする。このように設定した上で p および r が学習過程に及ぼす影響について考察を加える。

Fig. 4 には重み係数の初期値が 0 で、観測ノイズが混入しない場合について r の値を 1.0 に固定し、 p の値を $10, 10^2, 10^3, 10^4, 10^5$ としたときの重み係数マトリクスの (4, 2) 成分と、その重み係数の推定誤差分散に相当する P マトリクスの (17, 17) 成分の時間的変動の一例を示した。ただし、 Q_i については恒等的に 0 とした。重み係数の時刻歴より p の値が大きいくほど学習の進行速度が速くなることわかる。これは重みの初期値を 0 として学習を始めているため、学習の初期段階では重み係数の誤差が大きくなっておりその推定誤差分散も大きくとらなければならないことを表わしている。また P の値は式 (34) でカルマンゲイン、すなわち重みの修正量を決定する値であり、 p を大きくとることによって 1 回の修正量が大きくなるため収束が早くなっている。重み係数の初期値の与え方については後に詳しく述べるが、対象とするモデルについて先見的情報がある場合には、重み係数の初期値として、ある程度正確な値を与えることができる。このようなある程度正確な初期重みを与えて学習を始める場合は、初期段階での重み係数の誤差が小さいと考えられるので p の値をこれよりも小さくする必要はない。Fig. 4 に示すように、 P の要素の値は観測を重ねる度に小さくなっていくので、BP 法の場合と同じく学習の進行と共に修正速度が低下することは避けられない。

また、解析を行う上で、 p と r との比が一定であれば学習過程が同じになることがわかった。その例を Fig. 5 に示す。Fig. 5 の上段は p/r が 1 となるように p と r の値を設定した場合の重み係数マトリクスの (3, 3) 成分の時間的変化を示したものであり、以下同じく中段は p/r が 10、下段は p/r が 100 となるよう

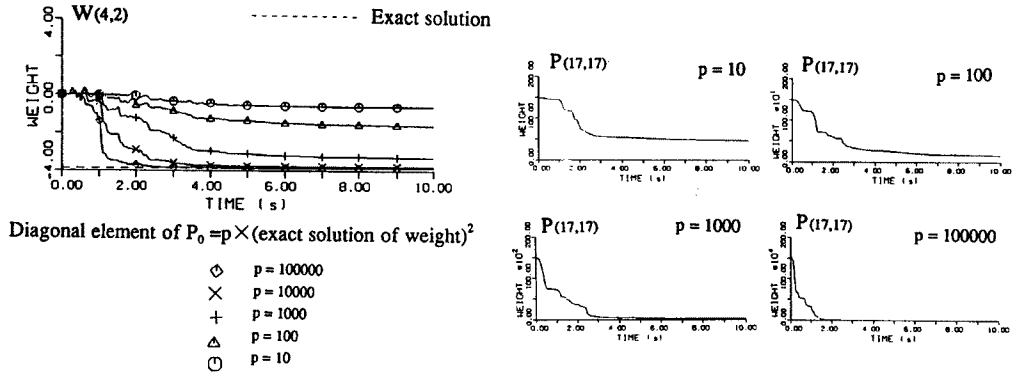


Fig. 4 The time history of weight $W(4,2)$ (left) and the time history of several element of the covariance matrix P (right).

に p と r を与えたものである。これにより、学習過程は p/r の値によって決定されていることがわかる。これは式 (34) でカルマンゲインを計算する時、その値が p と r の比によって決定されていることを表わしている。

観測データに混入するノイズについては後に述べるが、ノイズが混入する場合に、学習過程で r の値を変化させたときネットワークから出力される応答値と重み係数の (2, 2) 成分の時刻歴を Fig. 6 に示す。右側の図中の太実線は学習信号であり、細実線は教師信号である。左側の図中の実線は重みの同定過程であり、破線は厳密解である。これより適切な r の値を設定しなければ効率のよい学習が行われなことがわかる。これは、 R が観測データに含まれるノイズ成分の大きさを表わすパラメータであり、除去するノイズの量を決定しているため、ノイズのレベルを適切に評価する必要があることを表している。

5.3 ノイズに対する同定の安定性

学習に用いる観測データに混入するノイズが学習過程に与える影響について考察を加える。混入するノイズは Fig. 7 に示されている。応答速度の観測データに混入するノイズは周波数帯域 0 から 25Hz のピンクノイズに包絡線関数を乗じて作成した。応答変位に混入されるノイズは速度に混入するノイズの積分値とした。そして各階の 1 次モード形に比例するレベルのノイズが各階の観測値に混入されるとした。

混入されるノイズの最大レベルを構造物の示す最大応答値の 5% に設定し、ノイズが混入した観測データを用いて学習を行なった。そのときの重み係数マトリクスの時刻歴を (1, 3), (1, 4), (2, 3), (2, 4) の 4 つの要素を例にとり Fig. 8 に示す。比較のため厳密解を破線で示した。中央より左側の図が BP 法で、右側の図がカルマンフィルタ法によるものである。これより、BP 法では教師信号にノイズが混入する場合に学習が不安定になっているが、カルマンフィルタ法ではノイズの影響をうまく削除し、安定した学習を行っていることが分かる。

5.4 重み係数の初期値が同定結果に及ぼす影響

学習を行なうにあたり重み係数の初期値を設定しなければならないが、その値は同定の対象となる構造物についての先見情報を考慮に入れて設定する必要がある。重み係数の厳密な解がある程度わかる場合には、初期値としてその値を用いれば効率よく学習を行うことができる。以下では重み係数の初期値が学習過程に及ぼす影響について考察を加える。

ここでは重み係数の初期値を状態方程式を解いて求めた重み係数の厳密解に比例する形で与える。そして得られている先見情報の度合いをその比例定数で表現する。例えば、構造物について先見情報が全くないとするときには厳密解の 0%, ある程度厳密値が推定できる場合には厳密解の 80%, 90% といった値を初

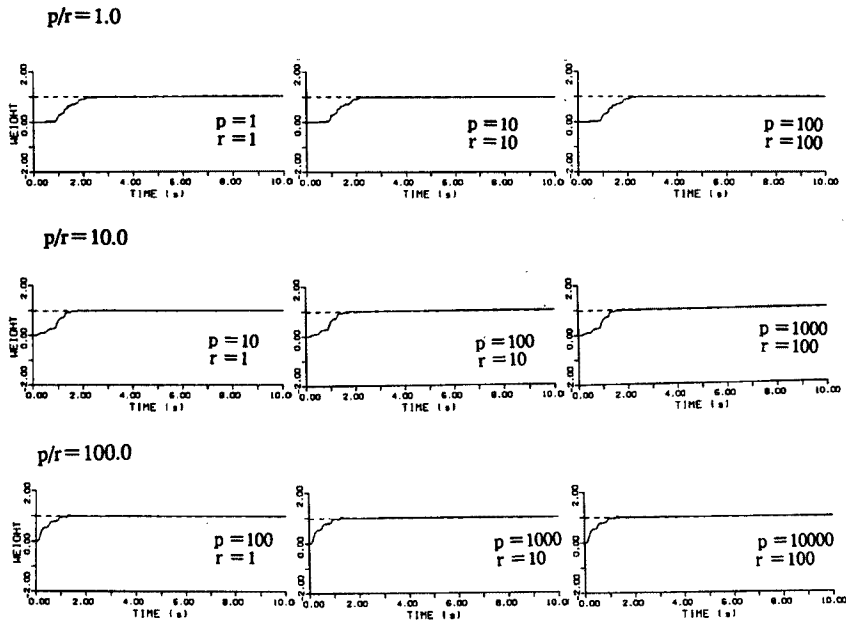


Fig. 5 Time history of weight $W(3,3)$ for several different ratios of p/r .

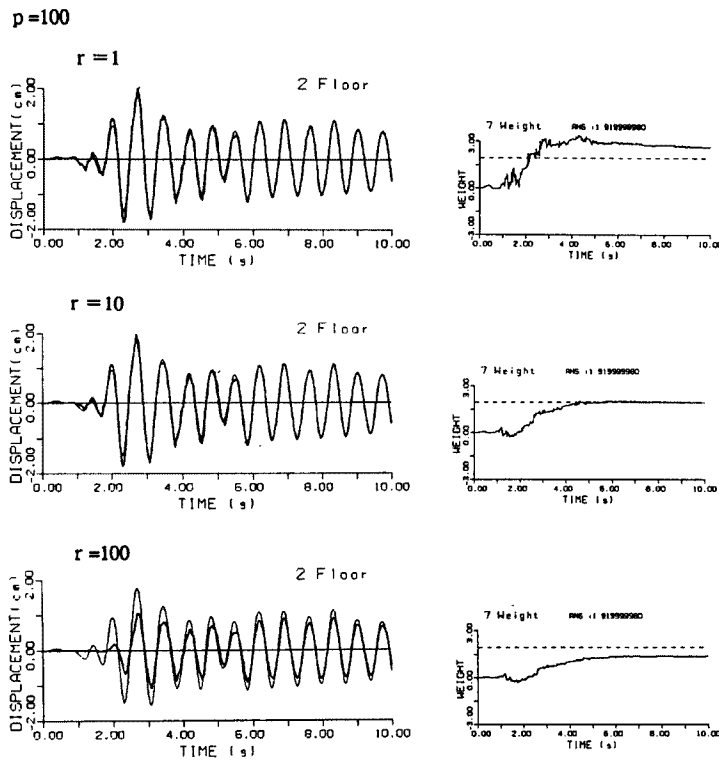


Fig. 6 The time history of structure response and weight $W(3,2)$ during a learning process (by Kalman filtering method).

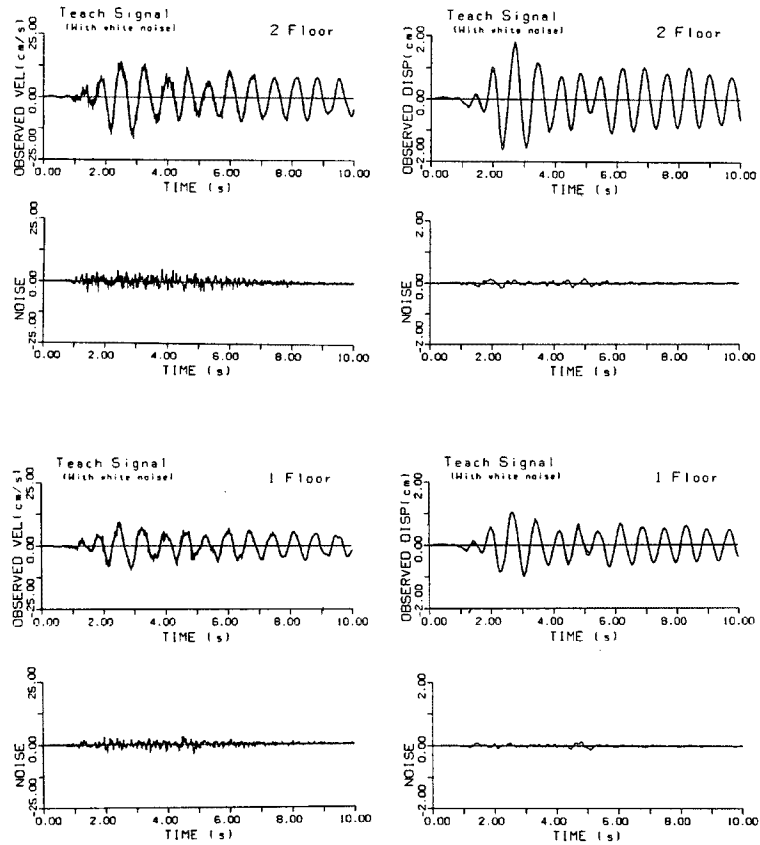


Fig. 7 Time history of noise contaminated teaching signal and noise.

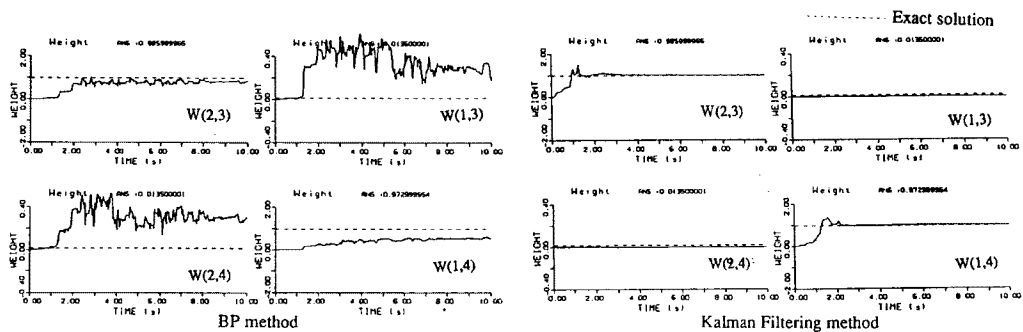


Fig. 8 Time history of weight (effect of noise).

期値として与える。

観測値にノイズが混入しない条件下で、BP法とカルマンフィルタ法の両方について、初期重みを厳密解の0%から110%まで変えて学習を行ない、重み係数の時間的変動の様子を重み係数マトリクスの(2, 1)成分を例にとってFig. 9に示す。これより、カルマンフィルタ法では初期重みの影響を受けず厳密解へと重み係数が収束しているが、BP法では重み係数の初期値により収束する値が大きく異なることが分かる。

そしてノイズが混入する場合も含め、それぞれの初期重みのケースについてグローバルな学習の繰り返しを10ステップ行なって、その結果最終的に得られた重み係数マトリクスの固有値計算を行ない、固有振動

数・減衰定数を求めた結果を **Table 2** に示す。

Table 2 によると BP 法を用いた場合には初期値として厳密解に近い値を選ばないと重み係数の収束状態が悪く、ノイズが混入しない理想的な場合においても、同定されるパラメータは厳密な値とは大きく異なっている。しかしカルマンフィルタ法では、ノイズが混入する場合においても初期値の設定に関わらず固有振動数と減衰定数がいずれも精度よく同定できていることが分かる。

10 秒間の学習を 1 ステップの学習とし、1 ステップの学習で最終的に得られた重み係数を次のステップの初期重みとして学習再度行うというグローバルな学習の繰り返しを行い、各ステップの学習が終わるごとに得られる重み係数に固定し、ネットワークにより応答計算を行った。

Fig. 10 には得られた応答計算結果の例を示す。ここに示したものはノイズがなく、重み係数の初期値を 0 として学習を行った場合である。カルマンフィルタ法では Fig. 5 に示した様に、重み係数が厳密解に収束しているので厳密な応答を再現できている。一方 BP による学習では数回のグローバルな学習を行うことにより、重み係数が厳密解とは大きく隔たった値に収束するにも関わらず、学習を重ねることで厳密な応答を再現することができた。これは BP 学習では局所解が得られていることを意味している。

このように BP 法により得られる重み係数が初期値の影響を受け、大きく異なるのは、重みの修正量が、

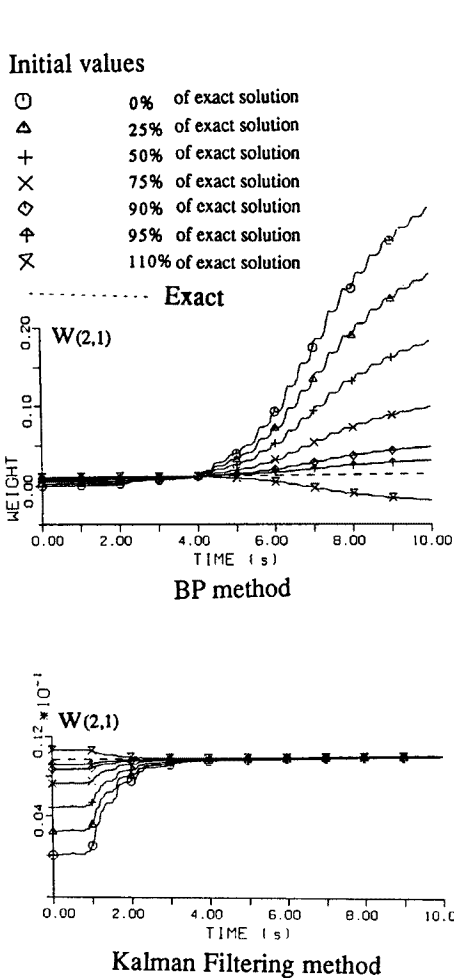


Fig. 9 Time history of weight (effect of initial values).

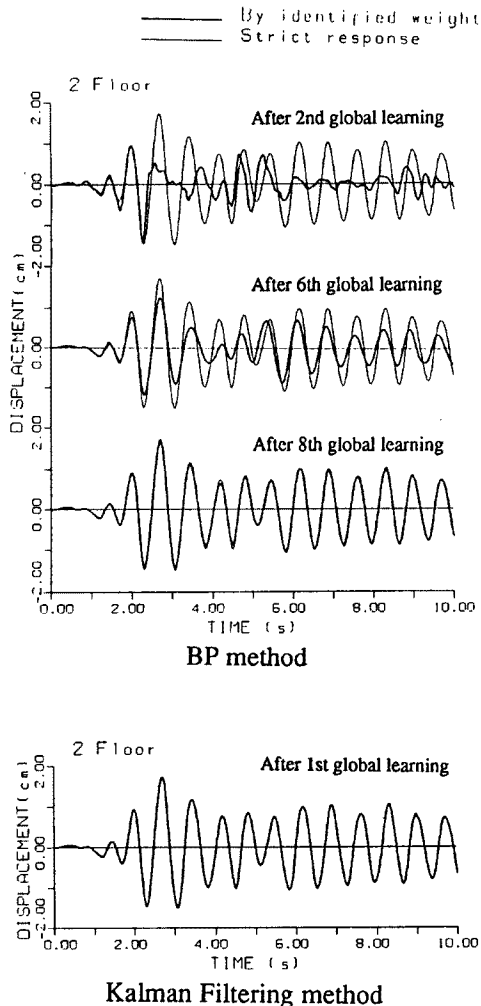


Fig. 10 Time history of response.

Table 2. Identified parameters.

		ω (rad/s)	h
strict	1st	8.63	0.00863
	2nd	22.6	0.0226

Initial weight		B. P. Algorithm				K. F. Algorithm			
		Without Noise		With Noise		Without Noise		With Noise	
		ω (rad/s)	h	ω (rad/s)	h	ω (rad/s)	h	ω (rad/s)	h
0%	1st	—	—	—	—	8.61	0.0120	7.98	0.1530
	2nd	—	—	—	—	19.2	0.0615	21.0	0.1050
50%	1st	8.65	0.0310	7.14	0.0839	8.64	0.0081	8.04	0.1300
	2nd	49.0	0.8460	39.7	0.9250	20.9	0.0166	22.7	0.0694
80%	1st	8.41	0.0076	7.24	0.0928	8.64	0.0097	8.09	0.1020
	2nd	30.3	0.4630	24.0	0.3640	21.9	0.0073	23.7	0.0469
90%	1st	8.43	0.0145	7.85	0.1480	8.64	0.0093	8.11	0.1060
	2nd	26.0	0.2640	21.7	0.0203	22.3	0.0153	24.1	0.0407
110%	1st	9.10	0.0029	8.90	0.1360	8.63	0.0068	8.13	0.1140
	2nd	20.2	0.2630	21.7	0.5180	22.9	0.0296	24.7	0.0275

ネットワークからの出力と観測値の差が最小になることだけを評価基準として、入力—出力の関係が教師信号と一致するように重みを決定するため、この条件を満たす重み係数の組み合わせが多く存在し、局所解に収束しているためであると考えられる。一方カルマンフィルタ法でこのように初期値にかかわらず重みが厳密解に収束することについては、さきに展開した重み係数の修正アルゴリズムの違いによるものと考えられる。すなわちBP法では重みの修正量を決定する際、出力層のk番目のノードに接続するリンクの重み係数の修正量の決定は、ただ出力層のk番目のノードにおける出力と観測値の誤差の情報のみによって行われ、他のノードにおける誤差は考慮されないのに対し、カルマンフィルタ法の学習アルゴリズムでは、全ての重み係数に相関があるとして推定誤差共分散行列Pを定義しており、重みの修正量が出力層の全てのノードにおける誤差を考慮に入れる形式で決定されているためであると考えられる。

6. 結論

本研究ではニューラルネットワークの学習過程に、拡張カルマンフィルタを適用することによって、新しい学習アルゴリズムを提案した。これを用いて構造物が地震時に示す挙動を再現するネットワークの構築を行った。以下に本研究で得られた成果を述べる。

- (1) ニューラルネットワーク理論により構造物の動的特性を支配するパラメータについての情報を予め与えることなく、地震時における構造物の応答の観測データとその時の地震観測記録の情報のみで応答特性を再現するネットワークを構築することができた。
- (2) カルマンフィルタをニューラルネットワークの重み係数の学習則として適用することにより、学習の際の教師信号である観測データにノイズが混入する場合においてもノイズの影響を削除し、安定した学習を行うことが可能になった。また、構造系の固有振動数と減衰定数を精度よく同定することができた。従来の誤差逆伝播法に比べカルマンフィルタを用いた学習法は、ニューラルネットワークの重み係数の学習法としてきわめて有効な手法であることを数値計算例を通して実証することができた。

以下には本研究に関する今後の課題を述べる。

まずカルマンフィルタを用いた重み係数の学習則では、ネットワークの構造が複雑になり、ネットワークを構成するリンクの数、すなわち学習しなければならない重み係数が多くなると、取り扱うマトリクスの次元が大きくなってしまいうえに計算時間が膨大になる。ゆえに取り扱うマトリクスの次元を小さくできるよ

に工夫することが必要になる。また学習を進行していく過程で、ノード間の結び付きが弱いリンク、すなわち重み係数が小さいリンクについてはそのリンクをカットできる忘却学習を行って、重み係数の最適解を学習するだけでなく、ネットワークの構造自体を学習し、最適なネットワーク構造を獲得するアルゴリズムを開発することが計算時間の短縮のみならず、システムの内部状態を明快に把握するためにも必要である。

参考文献

- 1) Kalman, R. E. : A new approach to linear filtering and prediction problems, Trans. ASME, J. Basic Eng., Vol. 82, pp.35~45, 1960.
- 2) Kalman, R. E. and R. S. Bucy : New result in linear filtering and prediction theory, Trans. ASME, J. Basic Eng., Vol. 83, pp.95~108, 1961.
- 3) 有本卓 : カルマン・フィルター, 産業図書, 1979.
- 4) 加藤寛一郎 : 最適制御入門, 東京大学出版会, 1987.
- 5) 片山徹 : 応用カルマン・フィルタ, 朝倉書店, 1983.
- 6) 星谷勝・斎藤悦郎 : データ解析と応用, 鹿島出版会, 1991.
- 7) 星谷勝・斎藤悦郎 : 拡張カルマン・フィルターを用いた同定問題の各種振動系への応用, 土木学会論文報告集, 第339号, pp.59~67, 1983.
- 8) 星谷勝・斎藤悦郎 : 線形多自由度系の動特性の推定, 土木学会論文集, 第334号/I-1, 289~298, 1984.
- 9) 土岐憲三・佐藤忠信・清野純史・市原和彦 : 地盤一構造系の非線形復元力の特性の同定, 京都大学防災研究所年報, No.31-B2, pp.1~21, 1988.
- 10) 計測と制御, 「システム同定特集号」, J. SICE, Vol.32, No.9, 1993.
- 11) Beck J. L. : Determining models of structures from earthquake records, California Institute of Technology, Pasadena, 1978.
- 12) Natke H. G. : Identification of vibration structures, Wien, New York, Springer, 1982.
- 13) Masri S F., Chassiako A. G. and Caughy T. K. : Structure-unknown non-linear dynamic system : identification through neural networks, Smart Materials and Structures, No. 1, pp.45-56, 1992.
- 14) 麻生英樹 : ニューラルネットワーク情報処理, 産業図書, 1988.
- 15) McCulloch, W. S., and Pitts, W. : A logical calculus of the ideas immanent in nervous activity, Bullet. Math. Biophysics 5, 115~133, 1943.
- 16) Rosenblatt. F. : The perceptron : A Probabilistic model for information storage and organization in the brain, Psychol. Rev. 65 [6], 386~408, 1958.
- 17) Rumelhart, D. E., Hinton. G. E., and Williams, R. J. : Learning representations by back-propagation errors, Nature 323-9, 533~536, 1986.
- 18) 石川真澄 : ニューラルネットワークの構造学習, 第17回システムシンポジウム, 第14回知能システムシンポジウム, 第1回ニューラルネットワークシンポジウム, 合同シンポジウム講演論文集, 特1~特6, 1992.