

構造化学習による非線形履歴振動系の同定

IDENTIFICATION OF HYSTERETIC VIBRATION SYSTEMS BY STRUCTURAL LEARNING

野田 茂¹⁾ 星谷 勝²⁾ 小淵康義³⁾

by Shigeru NODA, Masaru HOSHIYA and Yasuyoshi OBUCHI

This paper proposes a system identification technique to autoassociate hysteretic characteristics by learning a nonlinear dynamic structure which is represented by a system dynamics (SD) model. As this regard, neural network's learning serves as the means. The criterion function is defined by the sum of squared output errors which takes into account the energy of earthquake motion input to the structure. In order to obtain a regularly directed learning, effective use of the structural learning is considered as an alternative method for parameter and system identification problems. A global and local iteration procedure is proposed to obtain stable and fast convergency in the learning of recurrent networks.

Key Words: system identification, hysteretic structures, neural networks,
system dynamics model

1. はじめに

我が国のような地震国においては、構造物に作用する様々な動的外乱のうち、特に地震動の及ぼす影響が非常に大きい。構造物の振動解析においては、その終局状態を把握し、予測する上で、地盤・構造系への入力地震動とそれに対する応答を分析しなければならない。中でも、構造自体の性質を明らかにするシステム同定の研究は特に重要である。

システム同定とは逆問題の一つである。それは、ノイズに乱された実システムの入出力観測データに基づいて、与えられたクラスのモデルの中から、システムの構造つまり問題のシステムと等価な数学的モデルを求めることである。一般には、モデルのパラメーター、あるいはシステムの周波数応答やインパルス応答などを決定することが多い。そのためには、構造モデルの選択、システムの入出力データの測定、入力を選択やシステムとモデルの等価性を判定するための評価関数の設定などが重要となる¹⁾。

古典的な同定法には、周波数応答法、過渡応答法、相関法やスペクトル解析法がある。これは、正弦波、白色雑音や疑似白色雑音を入力として加え、周波数応答やインパルス応答を推定する方法である。

近代的な推定理論に基づいた同定法は、離散時間システムを対象にして、システムの支配方程式を既知とし、そのパラメーターを決定する方法である。これは、オフライン同定とオンライン同定の2つに分けられる。オフライン同定(バッチ処理)は、入出力の観測データを用いて同定する方法で、最小2乗法、一般化最小2乗法、補助変数法や最尤推定法などを含む。一方、オンライン同定法(シーケンシャル処理)は、データを得るごとに、パラメーターの推定値を更新していく逐次計算方式によって行われる。これには、逐次型最小2乗推定法、確率近似法や最尤推定法などが用いら

1) 工博 鳥取大学助教授 工学部社会開発システム工学科 (〒680 鳥取市湖山町南4-101)

2) Ph.D. 武蔵工業大学教授 工学部土木工学科 (〒158 東京都世田谷区玉堤1-28-1)

3) 東京電力(株) 葛野川水力建設所副長 第二工事事務所第二土木課 (〒409-06 山梨県大月市猿橋町殿上195)

れる。カルマンフィルタ（あるいは拡張カルマンフィルタ）を用いた同定法は逐次型最小2乗推定法である。

我々は、上記のような同定が「システム同定」と称するにふさわしいかどうか、疑問を投げかけたい。既往の同定法は、モデルの構造を先験的に既知と仮定しているため、結局はパラメータ同定法にはほかならないのである。すなわち、既往の方法は、パラメータ同定であるため、本来のシステム同定になっていないのである。物理的法則から、モデルの構造が既知で、パラメータのみが未知の場合にはそれでもよい。しかし、ブラックボックスと考えると、入出力データからモデルを構築する際には、発想の転換を図った「システム同定」を実施する必要がある。

振動・耐震の分野においては、これまで、地盤・構造系への入力と応答波形を分析することから、構造自体の性質を明らかにする同定の研究が行われている^{例えば 2)~5)}。しかし、既往の研究はパラメータ同定に限定されており、対象の構造や機能を把握する「システム同定」の研究は十分にされていない。そこで、本研究では、振動系を表すネットワークモデルを考え、予測のための基本ルールか先験情報が組み込まれたネットワークに対して構造化学習を行い、自己組織化によって、ダイナミクスモデルと振動特性を同定する手法を提案する。これには、ニューラルネットワークの学習機能を利用した情報処理を行う。

このニューラルネットワークを用いて履歴系の挙動を明らかにした研究としては、文献 6) や 7) が挙げられる。これらの研究では、1本の履歴ループの始点と終点を基本データとして、その間の曲線形を推定しており、いわゆる単純な関数近似問題を取り扱っている。しかし、基本的にはパラメータ同定に限定されている。しかも、その成果は既往の方法論^{2),3),5)}でも十分に説明のつく内容である。つまり、ニューラルネットワークを用いた動的システムの同定に関する既往の研究においては、新たな方向性や同定概念そのものに新規性がそれほど見られない。我々は、パラメータ同定を含む「システム同定」の本質を明らかにするため、ニューラルネットワークの考え方を活用することを試みる。従って、本研究では、その目的や枠組みが文献 6) や 7) のそれらと大いに異なっており、新たな特徴とユニークさを有している。

本研究では、履歴特性を有する非線形振動系を対象に、入出力の時系列観測データ（地震動と応答）を用いて構造化学習を行い、履歴特性と振動系のパラメータを同時に同定する方法を提案する。具体的には、エネルギー入力総量を導入するとともに、履歴振動系そのものの骨格構造を自己組織化するために、ニューロン間の重みに関する忘却項を付加した評価関数を考案し、グローバルかつローカルな繰返しによってモデルの表現能力を高める一連のアルゴリズムを示す。

2.では、システムダイナミクスの手法を用いて、対象とする非線形振動系モデルを多層のネットワークに表現する方法を説明する。本来、履歴振動系は未知であるが、ネットワーク構造を一般化させることは可能である。ここでは、同定の対象となる履歴系を先験的に既知とした場合を例にとり、Wen モデルとバイリニアモデルのニューラルネットワークを示す。3.ではエネルギー入力のバランス式を考慮し、システム同定のための評価関数の設定について、4.ではネットワークの骨格構造を自己組織化するための構造化学習について、5.ではシステム同定の学習アルゴリズムについて述べる。最後に、本研究のまとめを 6. に示す。

なお、本研究の概要については既に文献 8) と 9) に示してきた。これらの内容を取りまとめたのが本論文であることを断っておく。

2. 非線形履歴 1 自由度振動系の SD モデル

ニューラルネットワークを用いた構造化学習を実施する前に、振動系をネットワークで表現する。そのために、ここでは、レイト、レベル、定数や補助変数からなるシステムダイナミクス (SD) を用いる⁴⁾。フィードバックを含む時間的なダイナミクスをもったネットワーク (SD) を用意することにより、そのネットワークの特定部分のある時刻での状態量から、次時刻の状態量を表現することができる。

(1) 非線形履歴 1 自由度振動系

一般に、非線形履歴 1 自由度振動系の運動方程式 (変位 x 、速度 $\dot{x}$ と加速度 $\ddot{x}$) は、

$$\ddot{x} + 2\beta\omega_0\dot{x} + \omega_0^2g(x) = -\ddot{z}_g \quad \cdots \cdots (1)$$

で表せる。復元力特性 $g(x)$ は、履歴成分 z を用いると、次のようになる。

$$g(x) = \alpha x + (1 - \alpha)z \quad \dots\dots(2)$$

ここで、 β は減衰定数、 ω_0 は固有円振動数、 $\ddot{z}_g$ は入力加速度、 $\alpha(0 \leq \alpha \leq 1)$ は降伏比である。

復元力 $g(x)$ は x の履歴に依存する非線形な関数である。そのため、観測あるいは実験で得られた復元力-変位関係に基づき、多くの復元力モデルが提案されている。数値解析の立場から見ると、復元力モデルは、1) 直線の組み合わせからなるモデルと 2) 曲線によるモデルに分類される。

直線の組み合わせによるモデルの代表的なものとしてはバイリニアモデルが挙げられる。バイリニアモデルの履歴系の挙動は、

$$\dot{z} = [A - \gamma \text{sgn}(\dot{x}) \text{sgn}\{z - \bar{\beta} \text{sgn}(\dot{x})\}] \dot{x} \quad \dots\dots(3)$$

のように表せる。ただし、 A 、 $\bar{\beta}$ と $\gamma(A \geq \gamma \geq 0, \bar{\beta} \geq 0)$ はパラメーター、 $\text{sgn}(x)$ は符号関数 ($x \geq 0$ のときに $\text{sgn}(x) = 1$ 、 $x < 0$ のとき $\text{sgn}(x) = 0$) である。

一方、曲線によるモデルとして、Wen の多面型履歴復元力モデルが有名である。劣化しないときの履歴系の挙動は次式の微分方程式で表せる。復元力を柔軟にモデル化できるように、履歴曲線を制御する 4 つのパラメーター (A 、 $\bar{\beta}$ 、 γ と n) が用いられる。

$$\dot{z} = A\dot{x} - \gamma|\dot{x}|z|z|^{n-1} - \bar{\beta}\dot{x}|z|^n \quad \dots\dots(4)$$

(2) システムダイナミクスの適用

非線形履歴振動系を多層のネットワークに表現するため、ここではシステムダイナミクスの手法を用いることを提案する。

システムダイナミクスは、個々の制御とそれに基づく入出力関係を差分方程式で表現し、それらを連結して、対象全体の状態を表すものである。このシステムダイナミクスの特徴の一つとして、関数間に存在する非線形な関係を取り扱い易いということがある。

地震動と観測応答は既知であるが、それらを結びつける履歴モデル自体は、本来、先験的に与えられるものではない。非線形履歴振動系の同定の本質は、数式で容易に与えられることができない履歴モデルを見出すことにある。非線形履歴 1 自由度振動系をシステムダイナミクスで表示したニューラルネットワークのネットワーク構造は容易に一般化できる (後述の図 5 と図 6 参照)。しかし、ここでは、ニューラルネットワークの構造がどのように表されるのかを明示するため、簡単な数式で与えられる履歴モデル (先験的に既知) を対象に説明する。

まず、Wen の復元力モデル (式 (4) の $n = 2$ の場合) を例に、システムダイナミクスの手法を用いて、非線形履歴振動系をネットワークに展開する手順を示す。

時間刻み Δt が非常に小さいと、速度 $\dot{x}(t)$ は

$$\dot{x}(t + \Delta t) = \frac{x(t + \Delta t) - x(t)}{\Delta t} \quad \dots\dots(5)$$

の差分で表せるので、

$$x(t + \Delta t) = x(t) + \Delta t \dot{x}(t + \Delta t) \quad \dots\dots(6)$$

となる。同様に、

$$\begin{aligned} \dot{x}(t + \Delta t) &= \dot{x}(t) + \Delta t \ddot{x}(t + \Delta t) \\ &= \dot{x}(t) + \Delta t \{-2\beta\omega_0\dot{x}(t + \Delta t) - \alpha\omega_0^2x(t + \Delta t) \\ &\quad - (1 - \alpha)\omega_0^2z(t + \Delta t) - \ddot{z}_g(t + \Delta t)\} \quad \dots\dots(7) \end{aligned}$$

$$\begin{aligned} z(t + \Delta t) &= z(t) + \Delta t \dot{z}(t + \Delta t) \\ &= z(t) + \Delta t \{A\dot{x}(t + \Delta t) - \gamma|\dot{x}(t + \Delta t)z(t + \Delta t)|z(t + \Delta t) \\ &\quad - \bar{\beta}\dot{x}(t + \Delta t)z(t + \Delta t)^2\} \quad \dots\dots(8) \end{aligned}$$

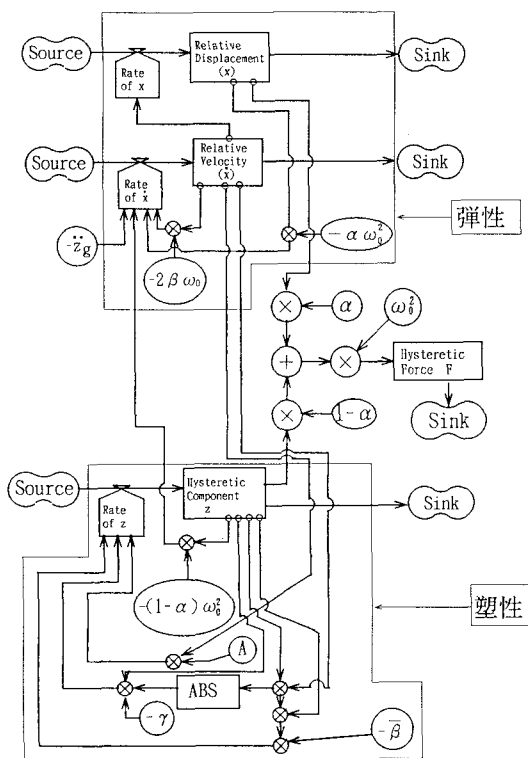


図1 Wenモデルによる履歴特性を示す1自由度振動系のフローダイアグラム

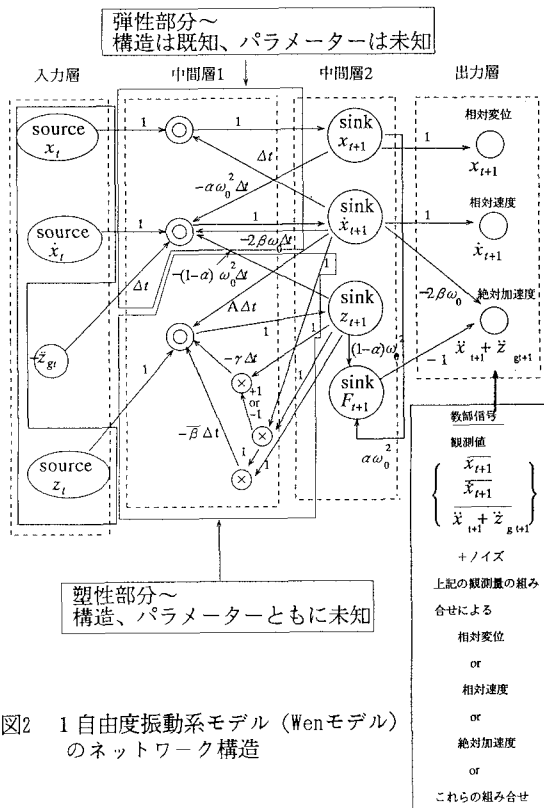


図2 1自由度振動系モデル (Wenモデル) のネットワーク構造

が導かれる。

これらの差分方程式は図1のようなフローダイアグラムで表せる。図において、Sourceは t 時点での状態量($x_t, \dot{x}_t$ と z_t)を、Sinkは $(t+1)$ 時点($(t+\Delta t)$ と同じ意味)での状態量($x_{t+1}, \dot{x}_{t+1}, z_{t+1}$ と復元力 F_{t+1})を意味する。図からわかるように、ネットワークは階層的であり、フィードバックを含むネットワークである。そのため、ある時刻での状態量から、次時刻の状態量を表現することができる。

図1のネットワークは、フィードバックを含む多層のネットワーク(図2)に変換できる。入力層は $x_t, \dot{x}_t, z_t$ と $-\ddot{z}_{g,t}$ ($-\ddot{z}_{g,t+1}$ ではなく、近似)のユニット、観測量は相対応答 $x_{t+1}, \dot{x}_{t+1}$ と絶対加速度($\ddot{x}_{t+1} + \ddot{z}_{g,t+1}$)のユニットからなる。履歴成分や復元力を直接的に計測することは困難なため、出力層としては、相対変位、相対速度と絶対加速度のいずれか、あるいはそれらの複数の状態量を考える。中間層1には図1のレート変数に対応するユニット($\odot$)と積の役割を果たすユニット($\otimes$)、中間層2には $x_{t+1}, \dot{x}_{t+1}, z_{t+1}$ と復元力 F_{t+1} を意味するユニットを含める。ニューラルネットワークのシナプス結合を意味する重み係数は、 ± 1 、 Δt や振動特性を表すパラメーター、すなわち、 $-2\beta\omega_0$ 、 $-2\beta\omega_0\Delta t$ 、 $\alpha\omega_0^2$ 、 $-\alpha\omega_0^2\Delta t$ 、 $(1-\alpha)\omega_0^2$ 、 $-(1-\alpha)\omega_0^2\Delta t$ 、 $A\Delta t$ 、 $-\gamma\Delta t$ と $-\beta\Delta t$ の値をとる。

履歴系の挙動がバイリニアモデル(式(3))で表せる場合も、同様の手順により、図3と図4のネットワークの関係が成立する。ただし、バイリニアモデルは、式(3)のように、sgn関数を含む。そのため、図4の中間層1には図3のsgnに対応するユニットを加えている。

本研究で提案するニューラルネットワークは、空間加算性、空間乗算性とsgn特性を考慮しており、入力重みの和、入力の積と重み和のsgn変換によって出力が結合されているという点で、既往のニューラルネットワークと異なる。なお、従来の研究ではニューロンすなわちユニットの入出力関係を非線形関数で表すことが多いが、ここでは線形関数を採用している。

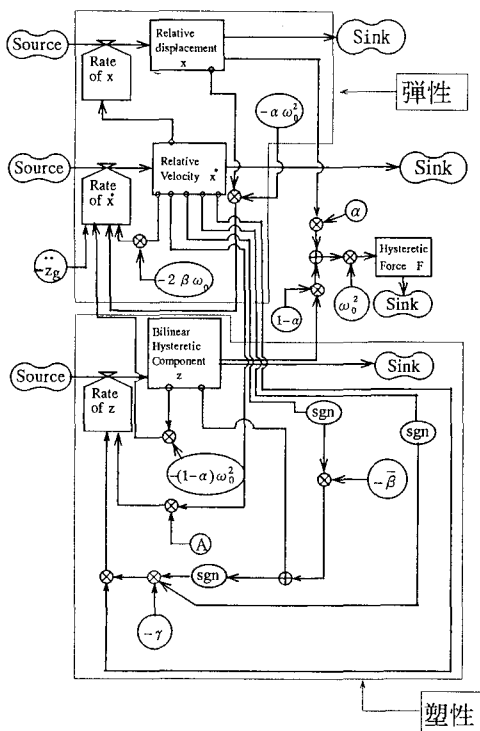


図3 バイリニアモデルによる履歴特性を示す1自由度振動系のフローダイアグラム

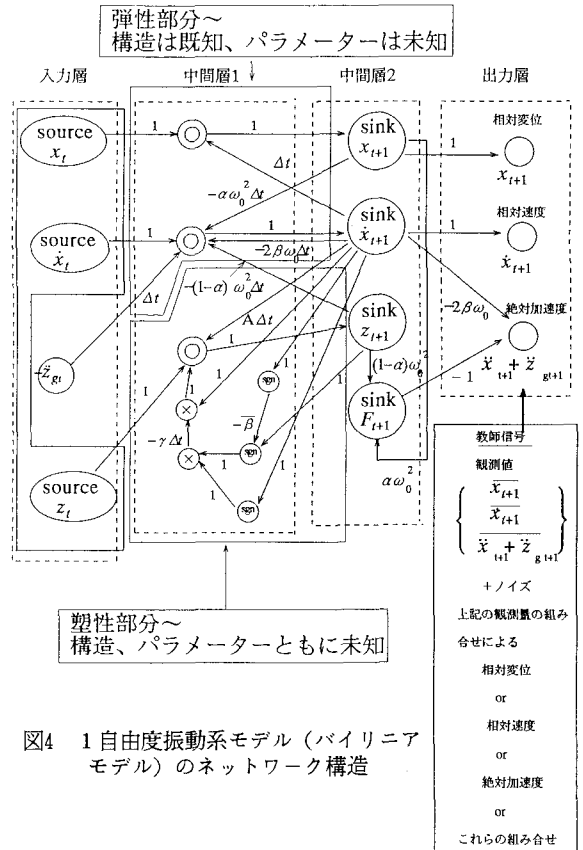


図4 1自由度振動系モデル(バイリニアモデル)のネットワーク構造

3. システム同定のための評価関数の設定

2. で述べた入出力層は観測量(入力加速度と応答)からなる。入力層には未観測の状態量の推定値が含まれている。このため、ニューラルネットワークの自己組織化では、観測情報をもとにして、正しい応答を示すように、そのネットワーク構造を変えていき、ニューロン間の結合度(すなわち、振動特性を表すパラメーター)を調整することができる。この実現のために、ここではバックプロパゲーション(BP)法¹⁰⁾を適用する。BPの目的は対象の特徴を把握することである。従って、観測値とネットワークの出力値の推定誤差を最小化する学習はその手段でしかない。しかし、このBP学習アルゴリズムでは、中間層数や各中間層でのニューロン数のようなニューラルネットワークのネットワーク構造パラメーターを事前に決定しておくことが必要である。従って、標準的なBP法では、最適なネットワーク構造を得ることができない。

本研究では、確率的BP学習と構造化学習アルゴリズムを適用する。これにより、ニューラルネットワークの入出力の観測値あるいは推定値を用いて学習することにより、ネットワークから不要なニューロンやシナプス結合を消滅させ、ネットワーク中に骨格構造を形成させることができる。

本章では、上記の学習を実施するため、システム同定における評価関数を新しく設定する。

(1) 推定誤差の最小化

図2と図4の多層ネットワークは、入力層と出力層の観測データ(相対変位、相対速度と絶対加速度)から、入出力間の写像関係を導き出すために考案したものである。この目的のため、ここではバックプロパゲーション学習法(以下BP学習法)を適用して、この写像関係を明らかにする。

BP学習法では観測量(目標出力、教師信号)の推定誤差最小化の原理を用いている。これは、ネットワークに入力データを与え、それに応じて計算された出力結果とその目標出力をできるだけ一致させるようにするものである。出力層の*i*番目のユニットの出力値を $F_{i,out}(t)$ 、教師信号(観測量)を $D_i(t)$ とすれば、推定誤差は $F_{i,out}(t) - D_i(t)$ で表せる。図2と図4からわかるように、各層のユ

ユニット間はリンクで結ばれている。そのリンク (シナプス) の重みは、次式の 2 乗推定誤差 $E(t)$ を最小にするように、求めればよい。

$$E(t) = \frac{1}{2} \sum_i (F_{i,out}(t) - D_i(t))^2 \quad \dots\dots\dots(9)$$

ここに、 i に関する総和は、全ての観測量について、計算することを意味する。

リンク重みとは図 2 と図 4 におけるパラメーターを意味する。例えば、図 2 においては、 $\pm 1, \Delta t, -2\beta\omega_0, -2\beta\omega_0\Delta t, \alpha\omega_0^2, (1-\alpha)\omega_0^2, -\alpha\omega_0^2\Delta t, -(1-\alpha)\omega_0^2\Delta t, A\Delta t, -\gamma\Delta t$ と $-\bar{\beta}\Delta t$ が対応する。リンク重みの修正は、基本的には、次式の最急降下法によって実施する。つまり、ユニット i とユニット j 間の重み w_{ij} の修正項 Δw_{ij} は、

$$\Delta w_{ij}(t) = -\eta \frac{\partial E(t)}{\partial w_{ij}} \quad \dots\dots\dots(10)$$

で表せる。重みは、 $w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t)$ ($w_{ij}(t)$ は学習回数すなわち時点 t のときの重み、 $t=0$ のときは初期値) に従って、修正される。ここに、 η は学習率である。重み w_{ij} が一定値に収束するまで、この操作は繰り返される。

(2) エネルギー入力項の導入

(1) においては、評価関数として、出力層における推定誤差 (式 (9)) を考えた。ここでは、さらに、拘束条件として、地震外乱による振動系へのエネルギー入力 の総量を導入する。そして、拘束条件を満たしながら、式 (9) の評価関数を最小化するために、拡張された評価関数を考案する。このようにしてエネルギー入力項を導入することはニューラルネットワークに物理的意味をもたせることに役立つ。すなわち、ニューラルネットワークに関する既往の学習法を用いれば、見かけ上、推定誤差は小さくなるものの、物理的に意味のある、あるいは真の重み (振動パラメーター) を必ずしも求められないのである。

まず、拘束条件を考える。そのため、式 (1) の両辺に $dx = \dot{x}d\xi$ をかけ、応答の継続時間 t_0 にわたってそれを積分すれば、次式が得られる。

$$\int_0^{t_0} \ddot{x}\dot{x}d\xi + \int_0^{t_0} 2\beta\omega_0\dot{x}^2d\xi + \int_0^{t_0} \omega_0^2g(x)\dot{x}d\xi = \int_0^{t_0} (-\ddot{z}_g)\dot{x}d\xi \quad \dots\dots\dots(11)$$

上式の右辺は、地震外乱による系へのエネルギー入力 の総量 (E_0) を表している。左辺の第 1 項は地震終了時の運動エネルギーを、第 2 項は消費エネルギー W_h を、第 3 項は塑性歪エネルギー W_p と地震終了時の弾性歪エネルギーを表す。地震終了時の運動エネルギーと弾性歪エネルギーは弾性振動エネルギー W_e を形成する。

その結果、式 (11) は、

$$W_p + W_e + W_h = E_0 \quad \dots\dots\dots(12)$$

のように書ける。

式 (11) の拘束条件を式 (9) に加えると、拡張された評価関数は、ラグランジェ乗数を λ とすれば、

$$\begin{aligned} E_f(t) &= \frac{1}{2} \sum_i (F_{i,out}(t) - D_i(t))^2 + \lambda \left\{ \int_0^t (-\ddot{z}_g)\dot{x}d\xi \right. \\ &\quad \left. - \left(\int_0^t \ddot{x}\dot{x}d\xi + \int_0^t 2\beta\omega_0\dot{x}^2d\xi + \int_0^t \omega_0^2g(x)\dot{x}d\xi \right) \right\} \\ &= \frac{1}{2} \sum_i (F_{i,out}(t) - D_i(t))^2 + E_n(t) \quad \dots\dots\dots(13) \end{aligned}$$

のようになる。

学習の際、ラグランジェ乗数 λ は、2 乗推定誤差項 (上式右辺第 1 項) に比して、大きめの値に設定しておく。これは、学習の初期段階において、2 乗推定誤差よりもむしろ、エネルギー入力項 (右

辺第2項)の拘束効果を勘案するためである。

学習を繰り返し行い、最適なネットワーク構造が得られると、つまりリンク重み(パラメーター)やユニット出力が最適値に近づくにつれて、式(13)のエネルギー入力項は次第に0に漸近していく。この場合、学習後の評価は、通常のBP法に比べて、それほど劣化しない。これにより、式(11)の拘束条件を満足し、かつ出力層の推定誤差を最小にするように、ネットワークは自動的に自己組織化する。

4. システム同定のための構造化学習

推定誤差の最小化からわかるように、BP学習法は入出力関係の再現を主目的としている。この場合、ネットワークをブラックボックスと考えている。そのため、中間層を解釈することが困難である。また、式(13)の評価関数を用いても、中間層の規則性を表現できる保証はない。このようなことから、学習データに内在する規則性をネットワーク上のリンク重み(すなわちネットワークの構造)あるいはユニット出力の規則性として出現させることが必要になる。

ここでは、上記の問題点を解決するために、構造化学習を取り入れる。

(1) 構造化学習アルゴリズム

構造化学習を用いれば、大きめのネットワーク構造から出発し、不要なユニットやリンクを逐次削除して、適切な大きさのネットワーク構造が求められる、つまり、これはネットワークに骨格構造を形成させる方法である¹¹⁾。この方法は次のような特徴を有する。第1に、ネットワークの大きさを予じめ決定する必要がなく、つまり中間層の数やユニット数の決定に対する試行錯誤が不要になる。第2に、構造の解釈が簡単になる。第3に、過小学習や過剰学習を防ぐことができ、一般化能力の向上が期待される。

そこで、ここでは、これらの利点を活かすように、システム同定を実施する。そのための考え方を以下に述べる。

a) 忘却付き学習

忘却付き学習では、BP学習による重み変化に加え、学習の繰り返し毎にリンク重みに関する忘却項を導入することによって、重みを減衰させる。これにより、不必要な小さいリンク重みは消去され、ネットワークの骨格構造を形成させることができる。また、忘却の効果によって、重みの初期値による影響も小さくなる。

学習時の評価関数としては、次式のように、BP学習による2乗推定誤差(式(13))にリンク重みの絶対値和の項を加えたものを考える。

$$E_b(t) = E_f(t) + \varepsilon' \sum_{ij} |w_{ij}(t)| \quad \dots\dots(14)$$

このとき、繰り返し毎の重みの変化量 $\Delta w_{ij}'$ は、

$$\Delta w_{ij}'(t) = -\eta \frac{\partial E_b(t)}{\partial w_{ij}} = \Delta w_{ij}(t) + \Delta w_{ij}^0(t) - \varepsilon \operatorname{sgn}_b(w_{ij}(t)) \quad \dots\dots(15)$$

で与えられる。ただし、 Δw_{ij}^0 とは式(13)の右辺第2項の偏微分(5章の式(24)に相当する)を意味する。 $\varepsilon(=\eta\varepsilon')$ は繰り返し毎の忘却量であり、 $\operatorname{sgn}_b(x)$ は、

$$\operatorname{sgn}_b(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases} \quad \dots\dots(16)$$

を意味する。

式(15)は、重み $w_{ij}(t)$ の値に拘らず、一定量 ε だけ変化することを示している。

b) 選択的忘却付き学習

忘却付き学習では、忘却項の導入によって、本来の重み変化 ($\Delta w_{ij}(t) + \Delta w_{ij}^0(t)$) から、常に ε だ

けずれる (式 (15)) ことになる。そのため、学習が進むと、学習到達度は劣化する。そこで、最適解に近ずいた場合、忘却は、不要と考えられる小さな重みに対してのみ、実施する。つまり、重みの絶対値 $|w_{ij}|$ が閾値 θ_{ij} 以下のときだけ、忘却付き学習を行うようにすればよい。

評価関数としては、式 (14) に比し、次式を用いる。

$$E_s(t) = E_f(t) + \varepsilon' \sum_{|w_{ij}(t)| < \theta_{ij}} |w_{ij}(t)| \quad \dots\dots\dots(17)$$

重みの変化量は次のようになる。

$$\Delta w_{ij}'(t) = \Delta w_{ij}(t) + \Delta w_{ij}^0(t) - \varepsilon \operatorname{sgn}_s(w_{ij}(t)) \quad \dots\dots\dots(18)$$

$$\operatorname{sgn}_s(x) = \begin{cases} 1, & 0 < x \leq \theta \\ 0, & x = 0 \\ -1, & -\theta \leq x < 0 \end{cases} \quad \dots\dots\dots(19)$$

この選択的忘却付き学習を用いれば、消滅した重みは復活することがないので、学習到達度は改善される。

(2) 確率的 BP 学習

従来の BP 学習では、1) 初期値を最小解付近に設定しない限り、極小解に陥るとそこにトラップする、および 2) 解が最適値に近づくにつれて、学習速度が遅くなるといった問題点を有している。これらは、リンク重みの修正に最急降下法を用いているために生じるものである。

本研究では、上記の問題点を克服するために、ノイズを付加する。その結果、重みの変化量は、式 (10) に対し、ランジェバン型微分方程式

$$\Delta w_{ij}(t) = -\eta \frac{\partial E(t)}{\partial w_{ij}} + \xi_{ij}(t) \quad \dots\dots\dots(20)$$

で表せる。

式 (20) において、 $\xi_{ij}(t)$ は、平均値 0 で、分散が $T(t)$ のホワイトノイズである。 $T(t)$ は、温度を意味し、学習の繰返し回数 t によって変化する。つまり、 $T(t)$ は、 t に依存して、次式のように、高温から低温に徐々に変化 (クーリングスケジュール、 T_0 = 初期値 (正の定数)) させる。

$$T(t) = \frac{T_0}{1+t} \quad \dots\dots\dots(21)$$

式 (20) により、学習の初期段階では、高温にすることで、極小解から脱出することができる。収束段階では、低温にすることによって、収束が妨げられないことになる。さらに、BP 学習では学習の収束性が悪いが、この方法を用いれば、安定した学習結果が得られ、学習の収束性の向上に役立つ。

5. 学習アルゴリズム

標準的な BP 学習アルゴリズムは一般化能力に欠ける。そのため、これでは、不規則に結合したニューラルネットワーク構造から出発して、不要なニューロンやシナプス結合を逐次削除し、適切なネットワーク構造が得られない。しかし、4. で述べた学習法とリカレントネットワークのための通時的 BP 学習法¹²⁾を組み合わせることにより、事前の構造情報なしに、フィードバックループを含む任意のネットワークを学習させることができる。すなわち、ニューラルネットワークの骨格構造は、最終的に、自己組織化する。この骨格構造により、入出力の関係を表す数理モデル (すなわち、式 (3) や式 (4) のような履歴系の挙動式) が構築できると同時に、構造の本質を解釈することが容易になる。つまり、我々の目的である「システム同定」は、試行錯誤を伴わずに、外乱と環境に応じて、自動的に実施できることになる。

本研究で提案したシステム同定の学習アルゴリズムは、通常の BP 学習法に比べ、次の点で異なる。

第 1 に、本研究では、物理的拘束を考えるために、評価関数にエネルギー入力項 (式 (11)) を付加している。第 2 は、図 2 と図 4 からわかるように、中間層 1 に 3 種類のユニットを考えていることである。つまり、1) 他ユニットの出力の加重和によってユニットへの入力を表す和ユニット、2) それを加重積で表す積ユニットと 3) 加重和を sgn 関数で変換した sgn ユニットのユニットを導入していることは、本研究の特徴の一つである。他の特徴としては、1) リンクの結合にフィードバックループを含むことと 2) ユニットの入出力間を線形関数でモデル化していることが挙げられる。

(1) エネルギー入力項の影響

式 (13) の拡張された評価関数のうち、エネルギー入力項 $E_n(t)$ を次式のように書き改める。なお、エネルギー入力項を導入すれば、ネットワークの学習に効果がある (3.2 参照)。

$$E_n(t) = \lambda \left\{ \int_0^t (-\ddot{z}_g) \dot{x} d\xi - \left(\int_0^t \ddot{x} \dot{x} d\xi + \int_0^t 2\beta\omega_0 \dot{x}^2 d\xi + \int_0^t \alpha\omega_0^2 x \dot{x} d\xi + \int_0^t (1-\alpha)\omega_0^2 z \dot{x} d\xi \right) \right\} \quad \dots\dots(22)$$

上式は、図 2 と図 4 におけるリンク重みやユニットを用いれば、次式のように表せる。

$$E_n(t) = \lambda \left\{ \int_0^t (-D_I(\xi)) F_{2,3}(\xi) d\xi - \int_0^t (F_{3,4}(\xi) - D_I(\xi)) F_{2,3}(\xi) d\xi + \int_0^t w_{2,3}^{3,4} (F_{2,3}(\xi))^2 d\xi - \int_0^t F_{4,3}(\xi) F_{2,3}(\xi) d\xi \right\} \quad \dots\dots(23)$$

ここに、 $F_{i,j}$ は j 層 i 番目のユニットの出力を、 w_{i,l_i}^{j,l_j} は l_i 層 i 番目のユニットから l_j 層 j 番目のユニットへのリンク重みを、 D_I は入力地震動を表す。ただし、上式において、 $\dot{x}$ は $F_{2,3}$ 、 $\ddot{x}$ は $(F_{3,4} - D_I)$ を意味するなどの対応関係を用いている。

式 (23) を重み w で偏微分したときに影響を受ける項は、 $w_{1,3}^{4,3}$ 、 $w_{3,3}^{4,3}$ と $w_{2,3}^{3,4}$ のリンク重みの項のみである。従って、次式のように、3 つの偏微分の式が重みの変化量 $\Delta w_{ij}^0(t)$ (式 (15) と式 (18) に対応する) に関係する。

$$\left. \begin{aligned} \frac{\partial E_n(t)}{\partial w_{1,3}^{4,3}} &= -\lambda \int_0^t F_{2,3}(\xi) F_{1,3}(\xi) d\xi \\ \frac{\partial E_n(t)}{\partial w_{3,3}^{4,3}} &= -\lambda \int_0^t F_{2,3}(\xi) F_{3,3}(\xi) d\xi \\ \frac{\partial E_n(t)}{\partial w_{2,3}^{3,4}} &= \lambda \int_0^t \{F_{2,3}(\xi)\}^2 d\xi \end{aligned} \right\} \quad \dots\dots(24)$$

(2) 中間層のユニットの分類

中間層 1 は 3 種類のユニットを含む。そのため、それぞれに対応して、偏微分の計算を実施しなければならない。

a) 和ユニットの場合

ユニットの入力項 $F_{i,l_i}(t)$ は、他ユニットの出力項 $F_{m,l_m}(t)$ と重み w_{m,l_m}^{i,l_i} を用いると、次式で表せる。

$$F_{i,l_i}(t) = \sum_{l_m=l_i-1}^{l_i+1} \sum_{m=1}^{\text{unit}(l_m)} w_{m,l_m}^{i,l_i} F_{m,l_m}(t) \quad \dots\dots(25)$$

ここに、 $\sum_{l_m=l_i-1}^{l_i+1}$ はフィードバックの影響を考慮した総和記号である。また、 $\text{unit}(l_m)$ とは、 l_m 層に

おけるユニット m の総数を意味する。

通常の BP 法と同様にして、式 (25) の偏微分をとると、次式が得られる。

$$\begin{aligned}\frac{\partial E(t)}{\partial w_{k,l_k}^{j,l_j}} &= \left\{ \sum_{l_n=l_j-1}^{l_j+1} \sum_{n=1}^{\text{unit}(l_n)} \frac{\partial E(t)}{\partial F_{n,l_n}(t)} \frac{\partial F_{n,l_n}(t)}{\partial F_{j,l_j}(t)} \right\} \frac{\partial F_{j,l_j}(t)}{\partial w_{k,l_k}^{j,l_j}} \\ &= \left\{ \sum_{l_n=l_j-1}^{l_j+1} \sum_{n=1}^{\text{unit}(l_n)} \frac{\partial E(t)}{\partial F_{n,l_n}(t)} w_{j,l_j}^{n,l_n} \right\} F_{k,l_k}(t) \quad \dots\dots\dots(26)\end{aligned}$$

b) 積ユニットの場合

ユニットの入力は加重積で表されるため、式 (25) と異なり、次式が得られる。

$$F_{i,l_i}(t) = \prod_{l_m=l_i-1}^{l_i+1} \prod_{m=1}^{\text{unit}(l_m)} w_{m,l_m}^{i,l_i} F_{m,l_m}(t) \quad \dots\dots\dots(27)$$

一方、偏微分は次式のようにになる。

$$\begin{aligned}\frac{\partial E(t)}{\partial w_{k,l_k}^{j,l_j}} &= \left\{ \sum_{l_n=l_j-1}^{l_j+1} \sum_{n=1}^{\text{unit}(l_n)} \frac{\partial E(t)}{\partial F_{n,l_n}(t)} \frac{\partial F_{n,l_n}(t)}{\partial F_{j,l_j}(t)} \right\} \frac{\partial F_{j,l_j}(t)}{\partial w_{k,l_k}^{j,l_j}} \\ &= \left[\sum_{l_n=l_j-1}^{l_j+1} \sum_{n=1}^{\text{unit}(l_n)} \frac{\partial E(t)}{\partial F_{n,l_n}(t)} \left\{ \prod_{\substack{l_m=l_n-1 \\ l_m \neq l_j}}^{l_n+1} \prod_{\substack{m=1 \\ m \neq j}}^{\text{unit}(l_m)} w_{m,l_m}^{n,l_n} F_{m,l_m}(t) \right\} w_{j,l_j}^{n,l_n} \right] \\ &\quad \times \left\{ \prod_{\substack{l_m=l_n-1 \\ l_m \neq l_k}}^{l_n+1} \prod_{\substack{m=1 \\ m \neq k}}^{\text{unit}(l_m)} w_{m,l_m}^{j,l_j} F_{m,l_m}(t) \right\} F_{k,l_k}(t) \quad \dots\dots\dots(28)\end{aligned}$$

c) sgn ユニットの場合

ユニットの入力は、次式のように、式 (25) で示した加重和を sgn 関数で変換したものである。

$$F_{i,l_i}(t) = \text{sgn} \left\{ \sum_{l_m=l_i-1}^{l_i+1} \sum_{m=1}^{\text{unit}(l_m)} w_{m,l_m}^{i,l_i} F_{m,l_m}(t) \right\} \quad \dots\dots\dots(29)$$

この偏微分をとると、次のようになる。

$$\begin{aligned}\frac{\partial E(t)}{\partial w_{k,l_k}^{j,l_j}} &= \left\{ \sum_{l_n=l_j-1}^{l_j+1} \sum_{n=1}^{\text{unit}(l_n)} \frac{\partial E(t)}{\partial F_{n,l_n}(t)} \frac{\partial F_{n,l_n}(t)}{\partial F_{j,l_j}(t)} \right\} \frac{\partial F_{j,l_j}(t)}{\partial w_{k,l_k}^{j,l_j}} \\ &= \left[\sum_{l_n=l_j-1}^{l_j+1} \sum_{n=1}^{\text{unit}(l_n)} \frac{\partial E(t)}{\partial F_{n,l_n}(t)} \right. \\ &\quad \times \delta \left\{ \sum_{l_m=l_n-1}^{l_n+1} \sum_{m=1}^{\text{unit}(l_m)} w_{m,l_m}^{n,l_n} F_{m,l_m}(t) \right\} w_{j,l_j}^{n,l_n} \left. \right] \\ &\quad \times \delta \left\{ \sum_{l_m=l_j-1}^{l_j+1} \sum_{m=1}^{\text{unit}(l_m)} w_{m,l_m}^{j,l_j} F_{m,l_m}(t) \right\} F_{k,l_k}(t) \quad \dots\dots\dots(30)\end{aligned}$$

ただし、

$$\delta(x) = \begin{cases} \infty & , x = 0 \\ 0 & , x \neq 0 \end{cases} \quad \dots\dots\dots(31)$$

(3) 学習の手順と結果のイメージ

以下の手順に従って、システム同定を実施する。

- 1) ある時間区間(時点 t と時点 $t+1$ の関係)において、1自由度振動系モデルのネットワークは図5で表せる。振動論的にみれば、ニューラルネットワークのネットワーク構造は弾性挙動と塑性挙動をする部分に分けられる。これに対し、初期のネットワーク構造と重み(すなわちパラメーター)をランダムに与える。すると、図5における塑性部分の初期状態のネットワークは、図6のようになる。図5において、構造が既知である弾性部分ではパラメーターのみを同定し、塑性部分においてはネットワーク構造とともにパラメーターをも同定する。
- 2) 拘束条件付きの拡張された評価関数に忘却項を追加して、BP学習アルゴリズムを構築する。なお、学習とともにノイズの分散を低減させるように、ランジェバン型微分方程式を与える。教師信号としては、観測あるいは実験データ、すなわち相対変位、相対速度や絶対加速度を用いる。なお、弾性挙動を示す中間層においては、1)で述べたように、ネットワークの構造つまりリンクの連結状態が既知なため、リンク重みすなわちパラメーターのみを同定する。
- 3) ネットワークの構造が明確になるように、選択的忘却付き学習アルゴリズムを導入し、学習を行う。ネットワークの構造がブラックボックスである塑性部分の中間層の構造の決定に当たっては、この学習法が威力を発揮することになる。
- 4) 2)～3)の方法により、観測データを順次取り組みながら、初期のネットワーク構造と重みを修正していくことにより、最適推定値が求められる。しかし、このままでは安定した解を得ることが困難である。そのため、図7のように、継続時間にわたって、一度同定された結果を初期条件として置き換え、再び構造化学習を行う。この考え方は著者の一人が提案したWGI法³⁾と同じである。ただし、各時間帯(時点 i と時点 $i+1$ の間、 $i=0 \sim n-1$)では、図8に示すようなローカルな繰返しを含む。図中の入、中と出は、ネットワークの入力層、中間層と出力層を意味している。このようなグローバル・ローカルな繰返しを何度か実施することにより、モデルの表現能力が高まる。

図5のネットワークはフィードバックを含む。そのため、上述のアイデアのみでは、通常のBP学習法のように、偏微分を正確に計算することができない。例えば、中間層1のユニット出力を計算するには、中間層2のユニット出力が必要となるが、その値は未知である。そのため、初期条件

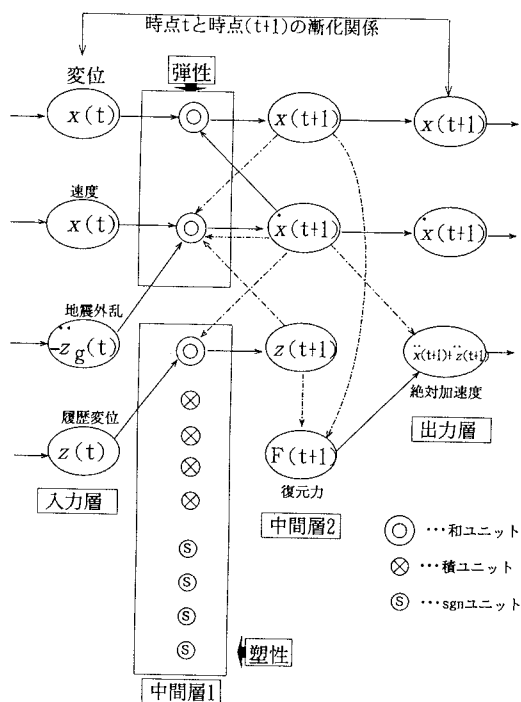


図5 1自由度振動系モデルのネットワーク

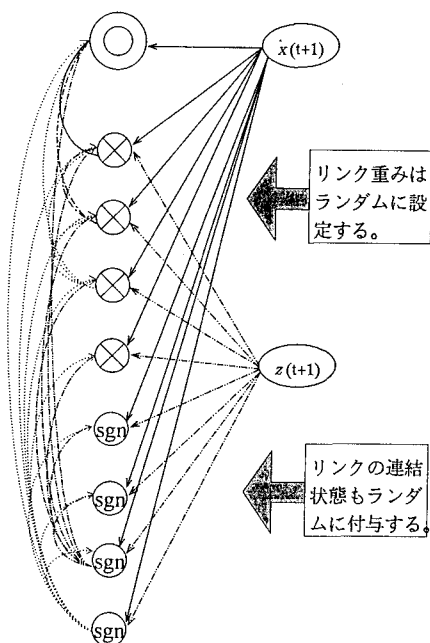


図6 塑性部分のネットワーク(初期状態)

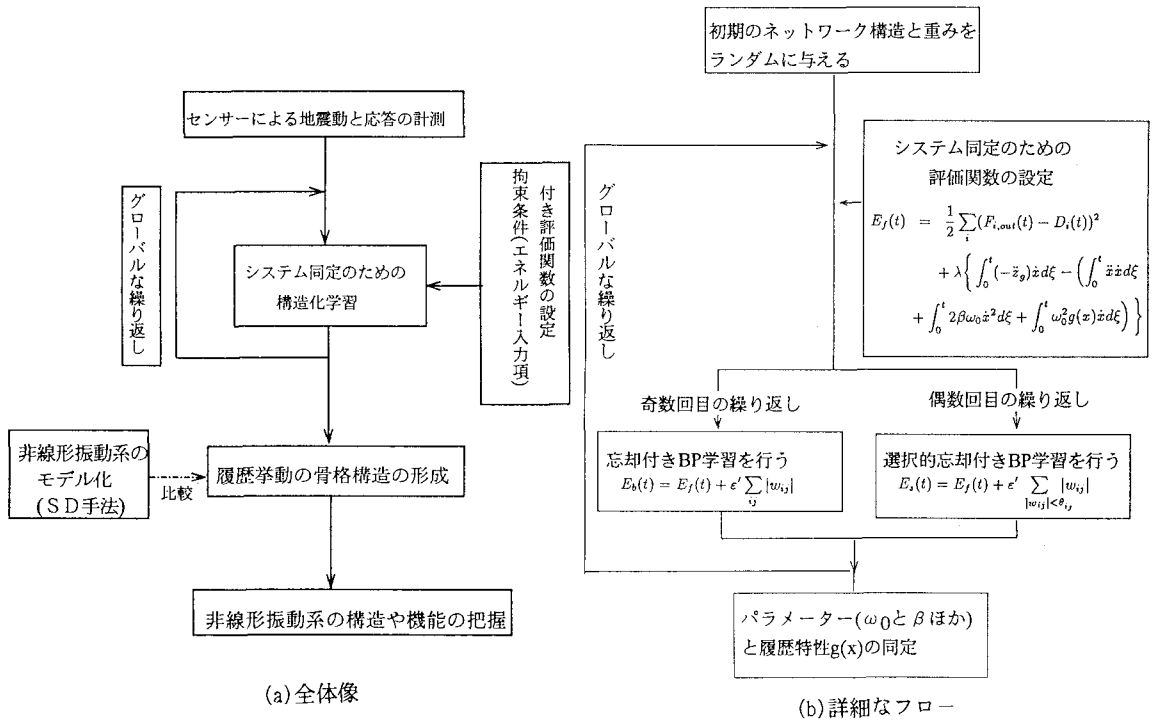


図7 システム同定のためのフロー

として、まず、中間層2のユニット出力は、入力層のユニット出力に等しいと近似する。しかし、図8に示すように、ローカルな繰り返しを実施する度に、その近似度は改善されるので、精度は高まる。偏微分を計算するときも、同様なテクニックを使う。

例えば、1940年 Imperial Valley 地震時に El Centro で得られた加速度記録 (NS 成分) に対応する応答 (相対変位、相対速度、相対加速度、相対変位と復元力の関係) は図9 (Wen モデルの場合) のようになる。手順1)において、初期状態における塑性部分のネットワークは図6のようである。しかし、手順4)を実施すると、図10や図11 (バイリニアモデルの場合) のように、履歴挙動の骨格構造が浮かび上がってくる。

ここでは、本論で提案した同定手法の概念を明示するため、簡単な数式で与えられる履歴モデルを対象に例示した。1.で述べたように、システム同定の本質は、本来、数式で容易に与えることができない観測 (あるいは実験) データ (入出力は既知であるが、内部は完全にブラックボックスである) を対象にすることにある。しかし、その場合でも、図5と図6のネットワークを用いて自己組織化を実施すれば、履歴振動系そのものの構造が自動的に同定されることは明らかである。

6. おわりに

本研究では、既発表の同定手法に対し、ニューラルネットワークの「関数学習機能」と「最適化機能」を利用した適応力のある同定方式を提案した。つまり、非線形履歴振動系を対象に、履歴系の挙動を表す数式やパラメータを同時に決定するためのシステム同定の考え方を提案した。

ここで述べたシステム同定の新しい枠組みは、自己組織化を用いて、内部構造が完全にブラックボックスのモデルを自動的に同定することにある。本論のアルゴリズムは一般的なシステム構造の同定を考えるものである。しかし、履歴モデルが先験的に未知であるという前提のもとで、どのようなニューラルネットワークの結合が得られるのかをイメージすることも重要である。そのため、ここでは Wen モデルとバイリニアモデルを例題にして、自己組織化の概念を示した。

システム同定は、具体的に、次の手順で行うことができる。

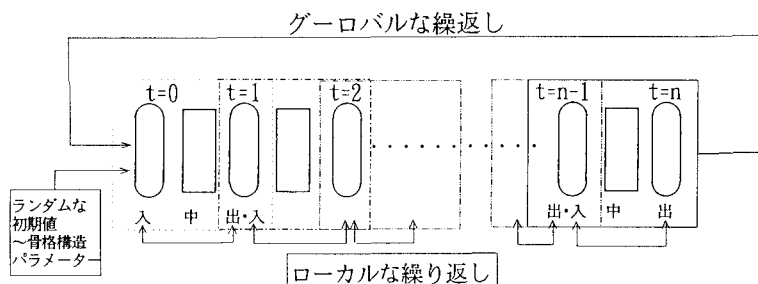


図8 グローバル・ローカルな繰返しを伴うネットワーク

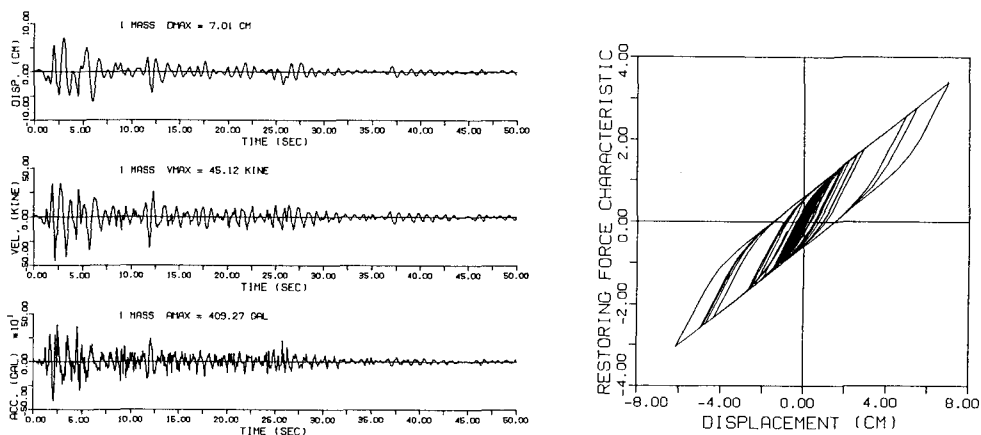


図9 Wenモデルによる履歴特性を示す1自由度振動系の応答特性の例

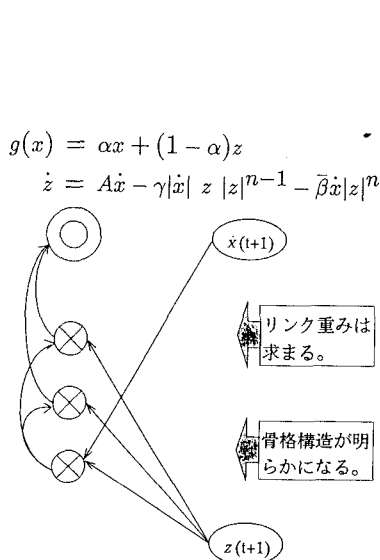


図10 Wenモデルによる塑性部分の骨格構造

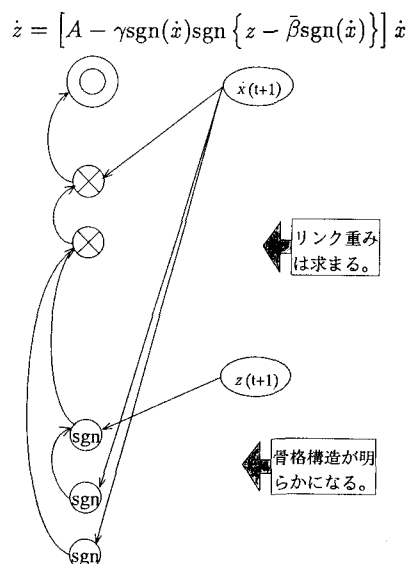


図11 バイリニアモデルによる塑性部分の骨格構造

- 1) 地盤・構造系への入力地震動とともに、構造系の主要な箇所の応答をセンサーで計測する。
- 2) システムダイナミクスの手法によって、非線形振動系をモデル化し、多層のネットワークに表現する。
- 3) 計測された入力と応答を教師信号とし、拘束条件(エネルギー入力項)付きの評価関数によっ

て構造化学習を行い、ネットワークに履歴挙動の骨格構造を形成させる。つまり、環境の変化に対して巧みに適応させ、かつネットワーク構造に対する学習の結果、対象の特徴抽出機能を自己組織化させる。

- 4) 3)の計算において、効率よく解が求められるように、継続時間にわたってはグローバルな、時間ステップではローカルな繰り返し同定を実施する。

以上の結果より、システムの入出力関係を表す履歴系の挙動は数式として表せることになる。本アルゴリズムによる同定の妥当性はシミュレーションによって検討することができる。すなわち、先験的に既知の履歴モデルに本方法を適用し、上記のようにして得られたネットワークの構造特性が2)のSDモデルと調和するかどうかを調べることで、有効性を判断することができる。

本方法によれば、試行錯誤を伴わず、外乱と環境に応じて、システム同定を自動的に実施できる。また、従来多用されてきたパラメーター同定に比し、提案した方式は、対象の構造や機能が把握できるため、汎用性があり、工学上の応用範囲が広いと言える。

本論文では「システム同定」の枠組みをとりまとめることを念頭においた。そのため、数値的かつ定量的な分析結果を十分に示していない。すなわち、著者らは、具体的な数値計算結果を呈示しなくても、同定概念そのものを見直して、物理的な意味があり、かつ精度の改善を図る方向性を述べることで、十分に価値のある内容を提案できたものと信じている。なお、具体的な数値計算例については別論文で報告する予定である。

参考文献

- 1) 相良 節夫・秋月 影雄・中溝 高好・片山 徹：システム同定，(社)計測自動制御学会，1987年11月。
- 2) 星谷 勝・丸山 収：非線形構造系の地震時挙動特性の同定，土木学会論文集，第386号/I-8，pp.397～405，1987年10月。
- 3) 星谷 勝・齊藤 悦郎：建設技術者のためのデータ解析と応用—カルマンフィルタを中心として，鹿島出版会，1991年7月。
- 4) Hoshiya,M., Obuchi,Y., and Noda,S.: Updating of dynamic structural systems by learnig, Probabilistic Mechanics and Geotechnical Reliability, A.S.C.E., pp.124～127, July 1992.
- 5) 野田 茂・久保田 秀樹：非線形履歴劣化多自由度系の震動特性の同定，第3回システム最適化に関するシンポジウム講演論文集，pp.195～209，1993年12月。
- 6) 山本 広祐：ニューラルネットワークによる履歴挙動のモデリングと数値解析への応用，構造工学論文集，Vol.38A，pp.85～94，1992年3月。
- 7) 山本 広祐・酒井 理哉：ニューラルネットワークによる実験履歴挙動の認識とネットワーク構築過程の考察，構造工学論文集，Vol.39A，pp.71～84，1993年3月。
- 8) 藤田 章弘・野田 茂・星谷 勝・小淵 康義：構造化学習による震動系のシステム同定と自己組織化，土木学会年次学術講演会講演概要集，第I部，pp.1336～1337，1992年9月。
- 9) 藤田 聡・野田 茂・星谷 勝・小淵 康義：構造化学習による非線形履歴振動系のシステム同定，土木学会第48回年次学術講演会講演概要集，第I部，pp.524～525，1993年9月。
- 10) Rumelhart,D.E., Hinton,G.E., and Williams,R.J.: Learning internal representations by error propagation, Parallel Distributed Processing—Explorations in the Microstructure of Cognition, Vol.1: Foundations (Edited by Rumelhart,D.E., McClelland,J.L., and the PDP Research Group), pp.318～362, 1989.
- 11) 石川 真澄：忘却を用いたコネクショニストモデルの構造学習アルゴリズム，人工知能学会誌，Vol.5, No.5, pp.595～603, 1990年9月。
- 12) 石川 真澄：コネクショニストモデルの忘却を用いた一般構造学習，電子情報通信学会技術研究報告，NC89-99, pp.55～60, 1990年。

(1993年9月16日受付)