

다층신경회로망을 이용한 동적인 시스템의 인식(Identification)모델에 관한 연구

*姜 東 佑, 李 尚 培

韓國海洋大學校 大學院 電子通信工學科

*Kang Dong Woo, Lee Sang Bae

Dept. of Electro. & Comm. Eng. Korea Maritime Univ.

ABSTRACT

In this paper, we introduce the identification model of dynamic system using the neural networks. We propose two identification models. The output of the parallel identification model is a linear combination of its past values as well as those of the input. The series-parallel model is a linear combination of the past values in the input and output of the plant. To generate stable adaptive laws, we prove that the series-parallel model is found to be preferable.

I. 서론

식별(Identification)문제는 직렬하게 매개변수화된 식별모델을 설정하는 것과 플랜트와 식별모델 출력사이의 오차를 최소화하기 위한 모델의 매개변수를 조절하는 것으로 구성된다. 플랜트의 표현에서 미선형 함수가 관심영역에 알려진 영역(Class)에 속한다고 가정되기 때문에, 식별(Identification)의 구조는 플랜트와 동일하게 선택된다. 가정에 의해 식별모델(Identification Model)에서 신경회로망의 연결강도 행렬(Weight Matrices)은 존재한다. 그래서 똑같은 초기조건에 대해서, 플랜트와 식별모델 둘 다 지정된 입력에 대한 똑같은 출력을 가진다.

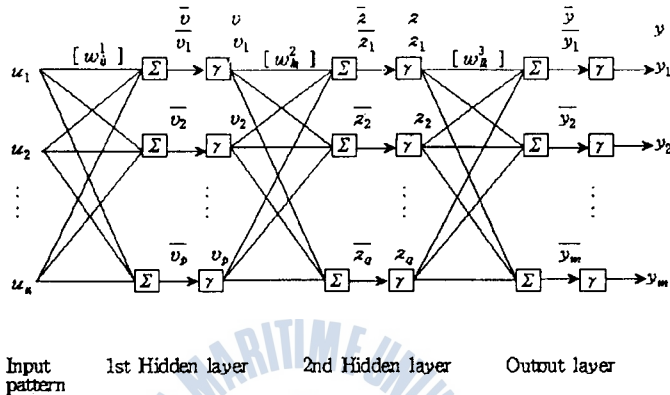
따라서 식별절차는 플랜트와 모델 출력사이의 오차에 기반을 둔 오차역전파(Error Back Propagation) 알고리즘을 사용하여 모델에서 신경회로망의 매개변수를 조절하는데 있다. 본 논문에서는 신경회로망을 이용하여 미선형 동적시스템(Dynamical System)에서 두 가지의 식별모델 (즉 **parallel identification model**, **series-parallel identification model**)을 비교·검토해 보았다. 본 논문의 II절에서는 신경회로망의 학습 알고리즘을 제시하였고 III절에서는 동적시스템의 구성을 나타내었으며 IV절에서는 병렬식별모델과 직렬식별모델의 구성을 보였으며 V절에는 두가지 제시한 식별모델(Identification Model)에 대해서 컴퓨터 시뮬레이션 결과를 제시하였다. 마지막으로 VI절은 결론이다.

-
- 1) 한국해양대학교 대학원 전자통신공학과
 - 2) 한국해양대학교 전자통신공학과 교수

II. 신경회로망의 학습 알고리즘

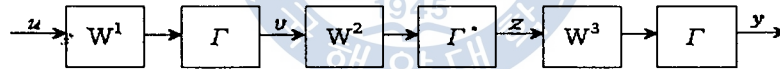
1.1 다층 신경회로망의 구조

Input Layer, Output Layer, 두 개의 Hidden Layer를 가진 전형적인 다층 신경회로망(Multilayer Network)는 그림 1에서 보여진다.



< 그림 1. 다층 신경회로망의 기본구조 >

그림 1을 다시 블럭도로 나타내면 그림 2와 같다.



< 그림 2. 다층 신경회로망의 블럭도 >

다층 신경회로망의 각 Layer와 입력·출력 Mapping의 함수는 다음과 같이 나타낼 수 있다.

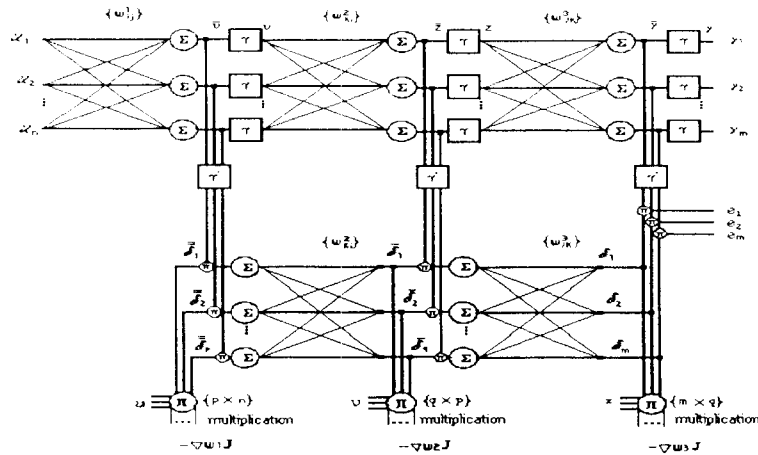
$$N_i [u] = \Gamma [W^i u] \tag{1}$$

$$y = N [u] = \Gamma [w^3 \Gamma [w^2 \Gamma [w^1 u]]] = N_3 N_2 N_1 [u] \tag{2}$$

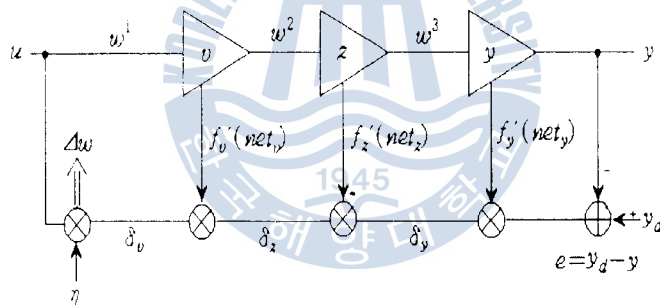
1.2 BACK PROPAGATION 학습

B-P 학습알고리즘의 기본원리는 입력층의 각 유니트(Unit)에 입력패턴을 주면, 이 신호는 각 유니트(Unit)에서 변화되어 중간층에 전달되고, 최후에 출력층에서 신호를 출력한다. 이 출력값과 목표값을 비교하여 차이를 줄여나가는 방향으로 연결강도(Weight)을 조절하고, 상위층에서 역전파하여 하위층에서는 다시 자기층의 연결강도(Weight)를 조정해 나간다. 실제 출력과 목표 사이에서는 모든 주어진 훈련 패턴에 대하여 최소화한다.

역전파(Back Propagation)을 위한 구조와 블럭도를 나타내면 그림 3, 그림 4와 같다.



< 그림 3. Back Propagation을 위한 구조도 >



< 그림 4. Back Propagation의 블럭도 >

여기서, 출력에서의 Squared Errors는

$$E = \frac{1}{2}(y_d - y)^2 \quad (3)$$

이고, y_d 는 바라는 출력이고 y 는 실제출력이다.

먼저, v neuron의 출력 $\rightarrow v$

z neuron의 출력 $\rightarrow z$

y neuron의 출력 $\rightarrow y$ 라고 각각을 가정하자.

출력 y 를 가지는 출력층(Output Layer)의 Error Signal Term는

$$\delta_y \cong - \frac{\partial E}{\partial (net_y)} = (y_d - y) \cdot f'_y(net_y) \quad (4)$$

출력 z 를 가지는 2번째 Hidden Layer의 Error Signal Term는

$$\delta_z \cong - \frac{\partial E}{\partial (net_z)} = - \frac{\partial E}{\partial z} \cdot \frac{\partial z}{\partial (net_z)} = (y_d - y) \cdot f'_z(net_z) \quad (5)$$

$$\delta_z = w^3 \cdot \delta_y \cdot f'_z(net_z) \quad (6)$$

출력 v 를 가지는 1번째 Hidden Layer Error Signal Term는

$$\delta_v \cong - \frac{\partial E}{\partial (net_v)} = - \frac{\partial E}{\partial v} \cdot \frac{\partial v}{\partial (net_v)} = (y_d - y) \cdot f'_v(net_v) \quad (7)$$

$$\delta_v = w^2 \cdot \delta_z \cdot f'_v(net_v) \quad (8)$$

따라서, update되는 weight는

$$w^1 \leftarrow w^1 + \Delta w^1 = w^1 + \eta \cdot \delta_v \cdot u \quad (9)$$

$$w^2 \leftarrow w^2 + \Delta w^2 = w^2 + \eta \cdot \delta_z \cdot v \quad (10)$$

$$w^3 \leftarrow w^3 + \Delta w^3 = w^3 + \eta \cdot \delta_y \cdot z \quad (11)$$

Ⅲ. 동적 시스템 구성

1.1 시스템의 입출력 표현식

동적인 시스템을 나타내는 미분방정식은 다음과 같다.

$$\frac{dx(t)}{dt} = \dot{x}(t) = \Phi [x(t), u(t)] \quad (12)$$

$$y(t) = \Psi [x(t)] \quad (13)$$

여기서, 상태변수 $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$, 입력 $u(t) = [u_1(t), u_2(t), \dots, u_p(t)]^T$, 출력 $y(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T$ 이다.

식 (12), (13)에서 Φ 와 Ψ 는 $\Phi: \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ 으로 정의되고 $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 으로 정의된다.

식 (12), (13)을 이산시간시스템에서 고려할 때 input-state-output 표현은 다음과 같다.

$$\begin{aligned} x(k+1) &= \Phi [x(k), u(k)] \\ y(k) &= \Psi [x(k)] \end{aligned} \quad (14)$$

(14)식에 의해 서술될 시스템이 선형이고 시불변으로 가정된다면 그것의 동작을 지시하는 방정식은 다음과 같이 표현될 수 있다.

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \quad (15)$$

시스템은 세쌍 $\{C, A, B\}$ 에 의해 매개변수화된다.

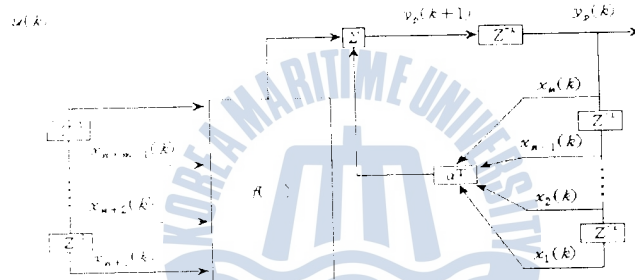
단일입력-단일출력(SISO) 플랜트에 대해서, (15)식에서 matrix A, B, C 는 플랜트방정식이 다음과 같이 쓰여질 수 있는 형태에서 선택되어질 수 있다.

$$y_p(k+1) = \sum_{j=0}^m a_j y_p(k-j) + \sum_{j=0}^m b_j u(k-j) \quad (16)$$

여기서, a_j, b_j 를 매개변수이다. 유석한 표현의 다입력다출력(MIMO)의 경우에 대해서 또한 가능하다. (16)식은 k 에서 출력이 입력과 출력 분타의 과거값의 선형조합이라는 것을 암시한다. 병렬식은 식별모델(Identification Model)의 선택에 동기를 부여한다.

$$y_p(k+1) = \sum_{j=0}^m a_j y_p(k-j) + f[u(k), u(k-1), \dots, u(k-m+1)] \quad (17)$$

여기서, $\{u(k), y_p(k)\}$ 는 k 에서 SISO 플랜트의 입·출력쌍을 나타내고, (17)식의 블록도는 그림 5에서 보여진다. (단, $m < n$)



< 그림 5. 단일입력단일출력(SISO) 플랜트의 표현 >

모델에서 $k+1$ 에서 플랜트의 출력은 입력 $u(k-j)$ ($j=0, 1, \dots, m-1$)의 과거 m 값 뿐만 아니라 그것의 $y_p(k-j)$ ($j=0, 1, \dots, n-1$)의 과거 n 값 둘 다에 의존한다.

함수 f 는 플랜트가 다중 신경회로망에 의해 나타내어지기 위해 관심의 영역에 알려진 영역에 속한다라고 믿리를 위해 가정할 것이다. 이 가정은 식별모델(Identification Model)의 선택에 동기를 부여하고 잘 드러난 식별문제의 진술을 허용한다. 특히 식별모델(Identification Model)은 플랜트와 같은 구조를 가지지만 조절할 수 있는 매개변수를 가진 신경회로망을 포함한다.

2) 시스템의 추정

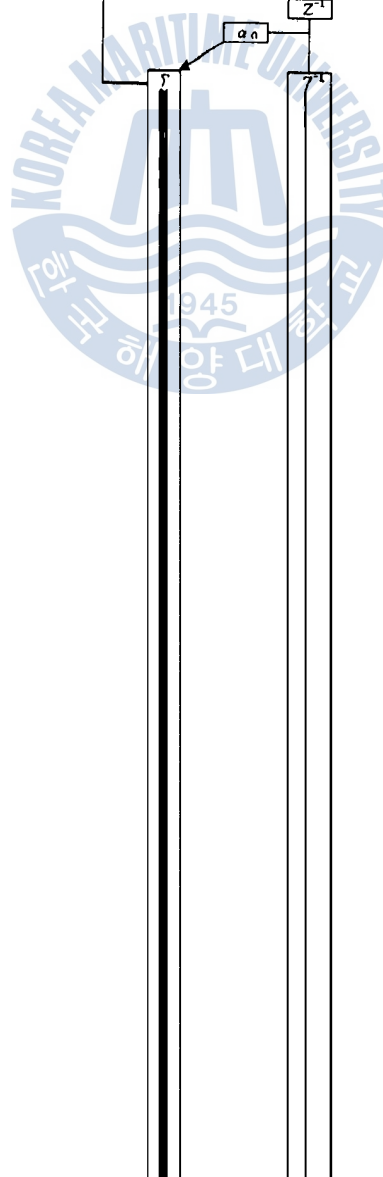
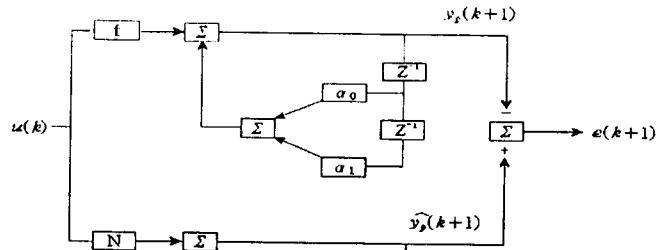
제어문제에 있어서 식별(Identification)모델을 설정하는 것이 중요한 과정이다. 측정된 신호로서 수학적인 모델의 범주에서 시스템의 동적특성을 결정하는 과정이므로 설정한 모델이 제어대상 시스템을 식별(Identification)할 수 있는 시스템 식별가능성(System Identifiability)이 있는 구조가 되어야 한다. 즉 주어진 모델 구조로써 대상시스템을 적절하게 나타낼 수 있어야 한다.

따라서, 본 논문에서 기존에 많이 응용하고 있는 병렬식별 모델(Parallel Identification Model) 보다는 직병렬식별 모델(Series-Parallel Identification Model)을 제안하고자 한다.

IV. 병렬식별 모델과 직병렬식별 모델의 구성

1.1 병렬식별 모델 (Parallel Identification Model)

이 모델은 우리가 일반적으로 사용하는 플랜트의 출력과 모델의 출력을 비교해서 오차를 최소화하는 방향으로 식별모델의 신경회로망을 학습하는 것이다. 그림 6은 $n=2, m=1$ 를 가진 모델에 의해 나타내질 수 있는 플랜트를 보인다.

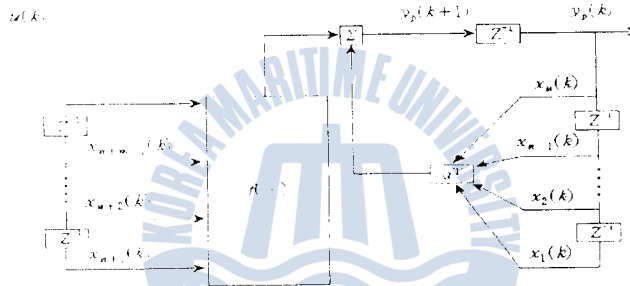


$$y_p(k+1) = \sum_{j=0}^n a_j y_p(k-j) + \sum_{j=0}^m b_j u(k-j) \quad (16)$$

여기서, a_j, b_j 는 매개변수이다. 유사한 표현이 다입력다출력(MIMO)의 경우에 대해서 또한 가능하다. (16)식의 $k+1$ 에서 출력에 입력과 출력 본나의 과거값의 선형조합이라는 것을 암시한다. 명칭식은 식별모델(Identification Model)의 선역에 놓이기 위하여한다.

$$y_p(k+1) = \sum_{j=0}^n a_j y(k-j) + f[u(k), u(k-1), \dots, u(k-m+1)] \quad (17)$$

여기서, $[u(k), y_p(k)]$ 는 k 에서 SISO 플랜트의 입·출력쌍을 나타내고, (17)식의 블록도는 그림 5에서 보이진다. (단, $m < n$)



< 그림 5. 단일입력단일출력(SISO) 플랜트의 표현 >

모델에서 $k+1$ 에서 플랜트의 출력은 입력 $u(k-j)$ ($j=0, 1, \dots, m-1$)의 과거 m 값 뿐만 아니라 그것의 $y_p(k-j)$ ($i=0, 1, \dots, n-1$)의 과거 n 값 둘 다에 의존한다.

함수 f 는 플랜트가 다중 신경회로망에 의해 나타내어지기 위해 관심의 영역에 알려진 영역에 속한다라고 믿리를 위해 가정할 것이다. 이 가정은 식별모델(Identification Model)의 선역에 동기를 부여하고 잘 드러난 식별문제의 진술을 허용한다. 특히 식별모델(Identification Model)은 플랜트와 같은 구조를 가지지만 조절할 수 있는 매개변수를 가진 신경회로망을 포함한다.

2) 시스템의 추정

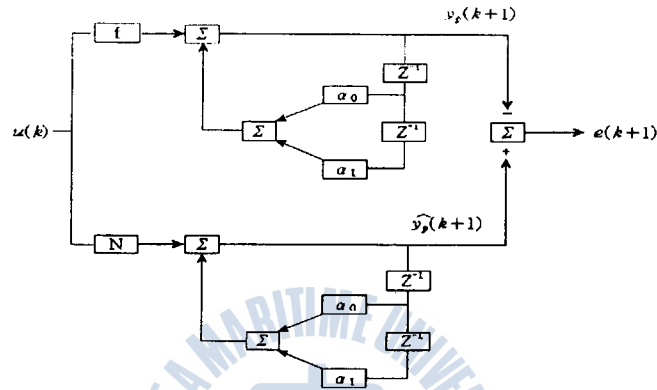
제어문제에 있어서 식별(Identification)모델을 설정하는 것이 중요한 과정이다. 측정된 신호로서 수학적인 모델의 범주에서 시스템의 동적특성을 결정하는 과정이므로 설정한 모델이 제어대상 시스템을 식별(Identification)할 수 있는 시스템 식별가능성(System Identifiability)이 있는 구조가 되어야 한다. 즉 주어진 모델 구조로써 대상시스템을 적절하게 나타낼 수 있어야 한다.

따라서, 본 논문에서 기존에 많이 응용하고 있는 병렬식별 모델(Parallel Identification Model) 보다는 직병렬식별 모델(Series-Paralle Identification Model)을 제안하고자 한다.

IV. 병렬식별 모델과 직병렬식별 모델의 구성

1.1 병렬식별 모델(Parallel Identification Model)

이 모델은 우리가 일반적으로 사용하는 플랜트의 출력과 모델의 출력을 비교해서 오차를 최소화하는 방향으로 식별모델의 신경회로망을 학습하는 것이다. 그림 6은 $n=2, m=1$ 를 가진 모델에 의해 나타내질 수 있는 플랜트를 보인다.



< 그림 6. 병렬식별 모델의 구성도 >

플랜트를 식별(Identify)하기 위하여 그림 6에 보여지는 식별모델의 구조를 가정할 수 있고 다음 방정식에 의해 서술된다.

$$\widehat{y}_p(k+1) = \widehat{\alpha}_0 \widehat{y}_p(k) + \widehat{\alpha}_1 \widehat{y}_p(k-1) + N[u(k)] \quad (18)$$

식별(Identification)은 모델출력 $\widehat{y}_p(k)$ 와 실제적인 출력 $y_p(k)$ 사이의 오차 $e(k)$ 에 기반을 둔 Back Propagation을 사용한다. 플랜트는 안정한 Bounded-Input Bounded-Output(BIBO)을 가진다고 가정한다. 따라서 플랜트에서 모든 신호(Signal)은 균등하게 나누어진다.

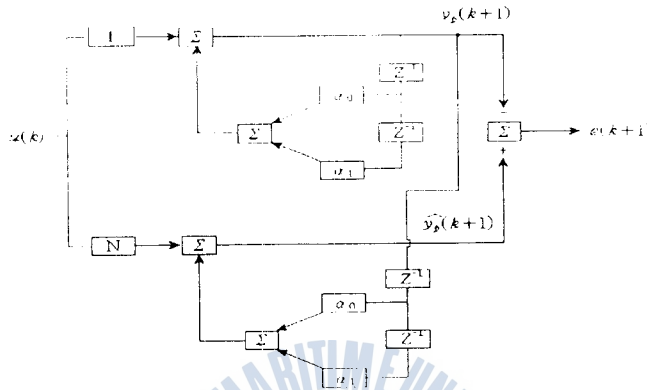
하지만 신경회로망과 함께 여기에 서술된 것으로서 식별모델의 안정성(Stability)는 확신될 수 없고 증명되어야 한다. 만약 병렬식별 모델이 사용된다면 매개변수가 수렴(Converge)하고 또는 그 출력오차가 영(Zero)가 될 것이라는 보장이 없다.

우리가 일반적으로 사용하는 병렬식별 모델보다는 다음에 제안하는 직병렬식별 모델은 플랜트의 출력이 다시 신경회로망으로 궤환(Feed Back)되어서 신경회로망에 영향을 미치므로 보다 효과적으로 플랜트를 식별할 수 있다.

1.2 Series-Parallel Identification Model

여기서는 병렬식별 모델보다 효과적으로 플랜트를 식별할 수 있는 직병렬식별 모델을 구성하였다. 그것은 그림 7에서 나타내어지며 플랜트의 직병렬식별 모델 출력에서 식별모델로 궤환된다는 것을 알 수 있다. 그래서 플랜트 출력의 지연(Delay)된 값들과 매개변수들과 곱해져서 신경회로망의 출력과 합해지므로 병렬식별 모델보다 효과적인 학습을 할 수 있다.

또, 플랜트가 인정한 BIBO이라고 가정되어 있기 때문에, 식별절차(즉, 신경회로망에 대한 입력)에서 사용된 모든 신호는 경계(Bound)된다. 게다가 플랜트의 입력과 출력의 과거값들은 신경회로망을 위해 입력벡터(Input Vector)를 형성하기 때문에 Static Back Propagation은 실질적으로 인산적 오버헤드(Overhead)를 감소시키는 매개변수를 조정하기 위해 이용될 수 있다.



< 그림 7. 직병렬식별 모델의 구성도 >

그림 7에서 보여 주는 식별 모델의 구조에 대한 방정식은 다음과 같이 서술된다.

$$\widehat{y}_p(k+1) = \widehat{a}_0 y_p(k) + \widehat{a}_1 y_p(k-1) + N[u(k)] \tag{19}$$

따라서, 여기서 제시한 직병렬식별 모델이 병렬식별 모델보다 효과적이라는 것을 보여주기 위해서 다음 절에서 여러가지 예제로서 시뮬레이션 결과를 보여준다.

V. 컴퓨터 시뮬레이션 결과

시뮬레이션 과정에서는, 제안한 모델을 사용하여 비선형 플랜트의 식별을 수행하기로 한다. 각각의 특성있는 4가지 예가 제시된다. 각 예제는 앞서 제시한 식별모델 즉, 직병렬식별 모델과 병렬식별 모델을 비선형 플랜트에 대해 비교하고, Back Propagation은 신경회로망의 매개변수를 조정하기 위해 사용되었다.

예제 1에서 4가지 플랜트방정식은

$$y_p(k+1) = 0.3y_p(k) + 0.6y_p(k-1) + f[u(k)] \text{ 이 사용되었고}$$

플랜트를 식별하기 위하여 직병렬식별 모델의 식별 모델식은

$$\widehat{y}_p(k+1) = 0.3y_p(k) + 0.6y_p(k-1) + N[u(k)] \text{ 이 사용되었고}$$

병렬식별 모델의 식별 모델식은

$$\widehat{y}_p(k+1) = 0.3\widehat{y}_p(k) + 0.6\widehat{y}_p(k-1) + N[u(k)] \text{ 이 사용되었다.}$$

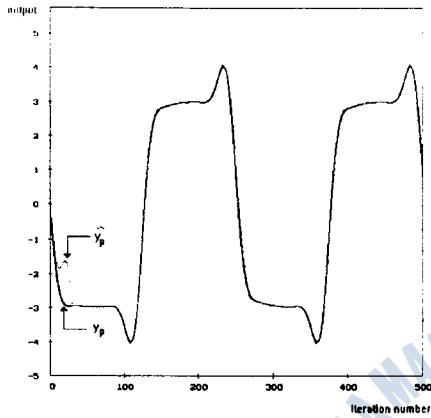
신경회로망의 구성은 입력이 1개, 첫번째 Hidden층이 20개의 Neuron, 두번째 Hidden층이 10개의 Neuron으로 구성했고, 출력은 1개이다. η 와 λ 는 0~1사이의 값을 사용했다.

<Example 1>

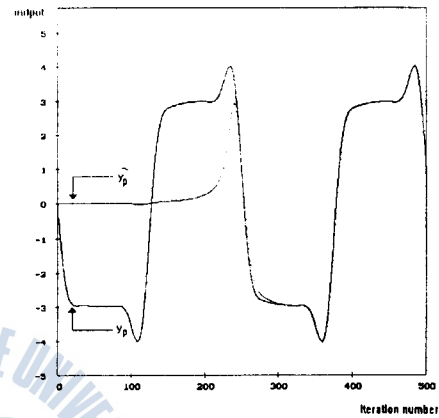
식별모델(Identification Model)에서 함수는

$$f(u) = 0.6 \sin(\pi u) + 0.3 \sin(3\pi u) + 0.1 \sin(5\pi u) \text{ 이고}$$

입력이 $s(k) = \sin(2\pi k/250)$, $u(k) = 0.5s(k) + 0.1 \cos(2\pi k/250)$ 일 때의 시뮬레이션 결과는 그림 8-1, 8-2와 같다.



<그림 8-1. 직병렬식별 모델 >



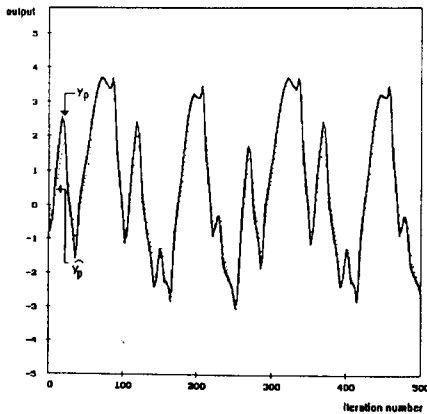
<그림 8-2. 병렬식별 모델 >

<Example 2>

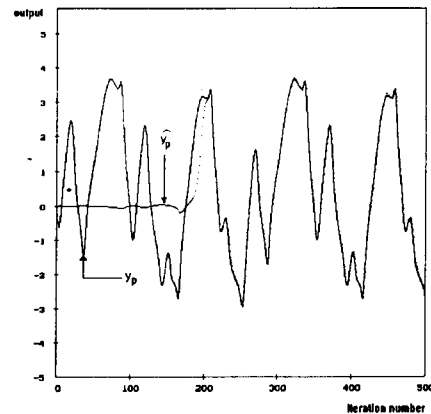
식별모델(Identification Model)에서 함수는 $f(u) = 0.6 \sin(\pi u) + 0.3 \sin(3\pi u) + 0.1 \sin(5\pi u)$ 이고

$$\text{입력이 } s(k) = \sin(2\pi k/250), \quad u(k) = \sin(\pi(s^2(k) + 0.3)) + 0.3 \sin(2\pi s(k))/(1 + s^2(k))$$

일 때의 좀 더 복잡한 비선형 시스템을 시뮬레이션한 결과는 그림 9-1, 9-2와 같다.



<그림 9-1. 직병렬식별 모델 >

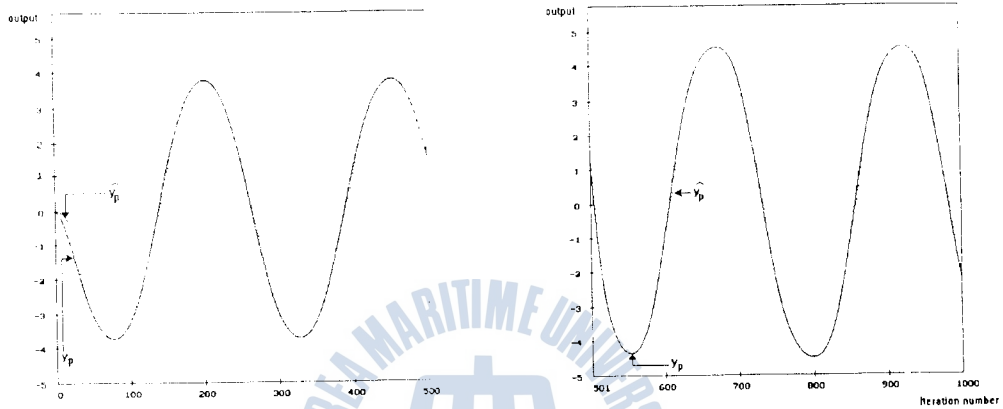


<그림 9-2. 병렬식별 모델 >

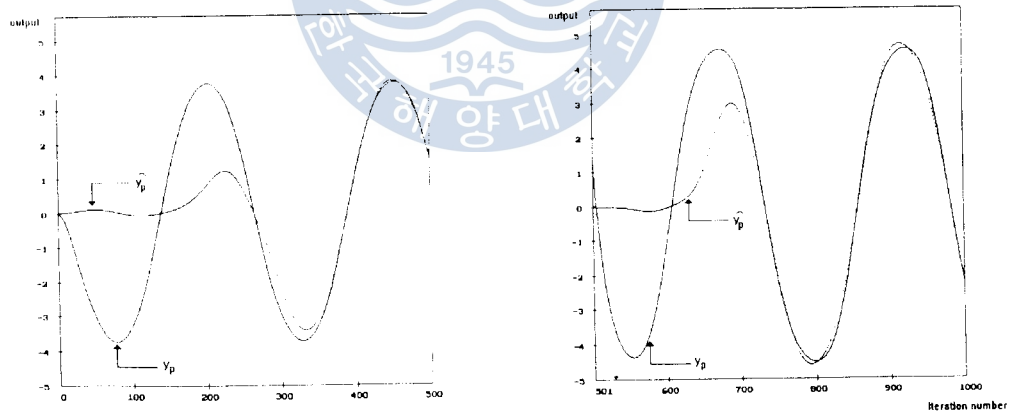
시스템에 주위의 외란이라든가 잡음이 유입되었을 때를 가상하여 도중에 입력을 변화시켜 시뮬레이션을 해보았다. 그 예는 예제 3, 4에서 나타내었다.

<Example 3>

식별모델(Identification Model)에서 함수는 $f(u) = 0.6\sin(\pi u) + 0.3\sin(3\pi u) + 0.1\sin(5\pi u)$ 이고
 Iteration $k=0$ 부터 500까지는 입력을 $u = 0.1\sin(2\pi k/250)$ 으로 주었고
 Iteration $k=501$ 부터 1000까지의 입력은 $u = 0.1\sin(2\pi k/250) + 0.1\cos(2\pi k/250)$
 와 같이 주었을 때, 시분레이션 결과는 그림 10-1, 10-2와 같다.



< 그림 10-1. 직방렬식별 모델 >



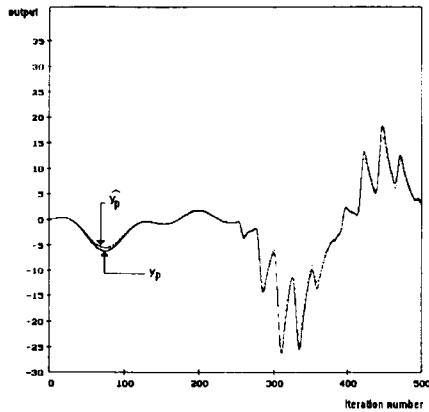
< 그림 10-2. 병렬식별 모델 >

<Example 4>

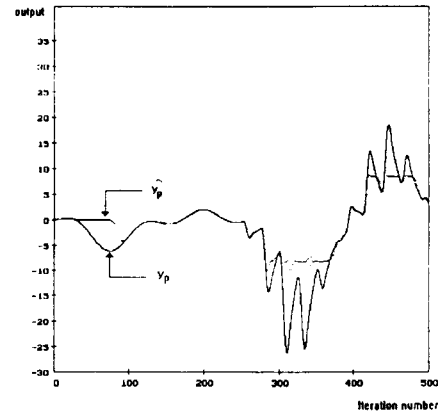
식별모델(Identification Model)에서 함수는 $f(u) = u^3 + 0.3u^2 - 0.4u$ 이고

$k=0 \sim 250$ 일 때는 $u(k) = \sin(2\pi k/250)$ 이고

$k=250 \sim 500$ 일 때는 $u(k) = \sin(2\pi k/250) + \sin(2\pi k/25)$ 으로 주었을 때 그림 11-1, 11-2와 같다.



< 그림 11-1. 직병렬식별 모델 >



< 그림 11-2. 병렬식별 모델 >

VI. 결론

본 논문에서 동적시스템(Dynamical System)의 식별모델(Identification Model)을 제시했다. 기존의 제어분야에서 많이 사용되고 있는 parallel identification model보다 플랜트의 출력이 식별 모델(Identification Model)로 궤환(Feed Back)되는 새로운 series-parallel identification model을 제안했으며, 시뮬레이션 과정을 통하여 제안한 방식인 series-parallel identification model이 parallel identification model보다 동적시스템에서 더 효과적이라는 것을 입증하였다.

앞으로 series-parallel identification model을 복잡한 Nonlinear Adaptive Control System에 이 용한다면 플랜트를 보다 효과적으로 제어할 수 있으리라 사료된다.

참고 문헌

- 1) K.S.Narendra and A.M.Annaswamy, *Stable Adaptive Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- 2) P.G.Gallman and K.S.Narendra. "Identification of nonlinear systems using a Uryson model," Becton Center, Yale University, New Haven, CT, tech. rep. CT-38, Apr. 1971; also *Automatica*, Nov. 1976.
- 3) L.Ljung and T.Soderstrom, *Theory and Practice of Recursive Identification*. Cambridge, MA: M.I.T.Press, 1985.
- 4) K.Hornik, M.Stinchcombe, and H.White, "Multilayer feed- forward networks are universal approximators," Dept. Economics, University of California, San Diego, CA, discussion pap., Dept. Economics, June 1988.
- 5) K.S.Narendra and K.Parthasarathy, "A diagrammatic representation of back propagation," Center for Syst. Sci., Dept. of Electrical Eng., Yale University, New Haven, CT, tech. rep. 8815, Aug. 1988.
- 6) G.C.Godwin, P.J.Ramadge, and P.E.Caines, "Discrete time multivariable adaptive control," *IEEE Trans.Automat.Contr.*, vol.25, pp.449-456, June 1980