

## 캐스케이드-상관 학습 알고리즘의 패밀리

# Family of Cascade-correlation Learning Algorithm

최명복 · 이상운

Myeong-Bok Choi and Sang-Un Lee

국립 원주대학 행정전산과 · 여성교양과

### 요 약

Fahlman과 Lebiere의 캐스케이드-상관 (CC) 학습 알고리즘은 신경망의 구성 알고리즘에서 가장 널리 사용되는 것 중의 하나이며, 망에서 은닉 뉴런을 캐스케이드 형태로 취함으로써 매우 강력한 비선형을 표현할 수 있다. 비록 이 역승이 유용할지 몰라도 대체로 문제를 푸는데는 강력한 비선형성이 요구되지 않으며 단점이 될 수도 있다. CC 알고리즘의 캐스케이드 구조 및 출력 뉴런의 가중치 훈련에 대한 변형된 형태인 3개 모델이 제안되고 경험적으로 비교되었다. 실험결과 다음과 같은 결론을 얻었다 : (1) 패턴분류에 있어서, 새로 추가되는 은닉 뉴런과 출력층간 연결강도만 훈련시키는 모델이 가장 좋은 예측력을 나타내었다; (2) 함수근사 문제에 있어서는 입력-출력 연결강도를 제거하고 시그모이드-선형 작동함수를 사용하는 모델이 CasCor 알고리즘보다 좋은 결과를 나타내었다.

### Abstract

The cascade-correlation (CC) learning algorithm of Fahlman and Lebiere is one of the most influential constructive algorithm in a neural network. Cascading the hidden neurons results in a network that can represent very strong nonlinearities. Although this power is in principle useful, it can be a disadvantage if such strong nonlinearity is not required to solve the problem. 3 models are presented and compared empirically. All of them are based on variants of the cascade architecture and output neurons weights training of the CC algorithm. Empirical results indicate the followings : (1) In the pattern classification, the model that train only new hidden neuron to output layer connection weights shows the best predictive ability; (2) In the function approximation, the model that removed input-output connection and used sigmoid-linear activation function is better predictability than CasCor algorithm.

**Key words** : 캐스케이드-상관 학습 알고리즘 (Cascade-correlation learning algorithm), 구성 알고리즘 (Constructive algorithm), 비선형성 (Nonlinearity), 작동함수 (Activation Function), 예측력 (Predictability)

## 1. 서 론

최근 들어, 신경망이 패턴분류와 함수근사 (회귀분석) 분야에 널리 이용되고 있다. 신경망을 훈련시켰을 때 좋은 일반화 성능을 얻기 위해서는 주어진 문제에 적합하도록 적절한 망의 구조 (즉, 모델)를 갖고 있어야만 한다. 너무 단순한 구조를 가진 망은 주어진 문제를 잘 학습할 능력이 없으며, 반면에 너무 복잡한 망은 훈련 데이터에 있는 잡음에도 적합되어 일반적으로 신경망의 일반화 성능이 저하되는 경향을 나타낸다[1]. 따라서, 신경망을 훈련시키는 과정에서 동시에 적절한 망의 구조도 자동적으로 찾아낼 수 있는 동적 알고리즘이 매우 요구되어진다[2]. 동적 알고리즘에는 전정 (Pruning 또는 Subtractive) 알고리즘과 구성 (Constructive 또는 Additive) 알고리즘 등이 있다[1-3]. 전정 알고리즘은 필요한 망의 구조보다 큰 망으로 초기화하여 원하는 해를 얻을 때까지 불필요한 뉴런이나 가중치를 제거하는 방법으로 훈련을 시작하기 전에 주어진 문제에 적합한 충분한 망의 크기를 얼마로 결정할 것인가를 판단할 수 없는 점이 가장 큰 단점이다. 이에 반해 구성 알고리즘은 작은 망으로 시작하여 만족한 해를 얻을 때까지 은닉 뉴런과 연결 가중치를 추가하

는 방법으로 초기에 매우 작은 은닉 뉴런 (은닉 뉴런이 없는 상태도 가능)으로 시작할 수 있기 때문에, 초기 은닉 뉴런 수를 즉시 결정할 수 있다. 주어진 문제에 적합한 망의 구조는 다양하게 있을 수 있으나, 구성 알고리즘은 전정 알고리즘에 비해 항상 작은 구조를 가진 해를 찾는 장점이 있다. 이와 같은 이유로 인해 신경망 분야에서 망의 구조를 자동적으로 결정하는 기법으로 구성 알고리즘을 더 선호하는 경향이 있다[3].

구성 알고리즘에 대해 다양한 연구가 있었으나, 가장 보편적으로 적용되는 알고리즘 중의 하나가 캐스케이드-상관 학습 알고리즘 (Cascade-Correlation Learning Algorithm : CC Algorithm)이다[1]. CC 알고리즘은 입력과 출력만으로 구성된 망 (초기 은닉 뉴런 수 = 0)에서 시작하여 망의 출력 오차와 망에 추가될 후보 뉴런의 출력과의 상관 (엄밀히 말해 공분산)이 크도록 후보 뉴런을 훈련시키고 후보 뉴런 들 중 가장 상관계수가 큰 후보 뉴런을 기존의 은닉층에 캐스케이드 형태로 추가시키는 방법으로 Fahlman과 Leibeire[4]는 은닉층-출력층 뉴런의 작동함수로 시그모이드-시그모이드 함수를 사용하였다. 이 방법은 원하는 해를 얻을 때까지 반복하여 망의 구조를 결정한다. CC 알고리즘에 대해 다양한 변형 형태가 연구되었으나 CC 알고리즘보다 월등히 좋은 결과를 얻지는 못하였다[5].

본 논문은 CC 알고리즘에 기반을 두고 CC 알고리즘의 구조와 훈련 방법을 변형시킨 3개의 모델을 제시하고자 한다. 2장에서는 CC 알고리즘 관련 연구의 문제점을 살펴본다. 3

접수일자 : 2004년 6월 8일

완료일자 : 2004년 8월 3일

장에서는 CC 알고리즘을 변형시킨 3개의 모델을 제시하고, 4장에서는 실험을 통해 변형된 CC 알고리즘이 패턴분류나 함수근사 문제에 보다 적합함을 보인다.

## 2. CC 알고리즘 관련 연구

### 2.1 CC 알고리즘

CC 알고리즘은 2가지 주요 아이디어를 내포하고 있다. 즉, 후보 뉴런을 상관(공분산)이 최대가 되도록 훈련시키며, 후보 뉴런이 은닉층에 추가되면서 은닉층이 캐스케이드 형태를 취하도록 하는 기법으로 새로 추가되는 은닉 뉴런은 신경망 입력과 기존에 추가된 은닉 뉴런들 모두에 연결이 된 형태로 새로운 1개의 은닉층을 형성하여 복잡한 캐스케이드의 다층 구조를 형성한다. 모든 훈련 데이터에 대해 뉴런의 출력  $H_i$  를 가진 후보 뉴런과 출력 뉴런  $o$ 의 출력 오차  $E_o$ 에 대한 공분산  $S$ 는 식 (1)로 구해진다[4].

$$S = \sum_0 \left| \sum_i (H_i - \bar{H})(E_{i,o} - \bar{E}_o) \right| \quad (1)$$

여기서,  $\bar{H}$ 는 모든 입력 데이터에 대한 후보 뉴런의 평균 출력이며,  $\bar{E}_o$ 는 출력 뉴런  $o$ 의 평균 오차(즉, 실제 출력 값과 목표 출력 값과의 차이)이다. 따라서 CC 알고리즘의 핵심은 식 (1)과 같이 신경망 출력의 오차와 망에 추가될 후보 뉴런 출력 사이의 공분산이 최대가 되도록 망을 훈련(후보 뉴런의 입력 가중치 조절)시키고 후보 뉴런들 중 공분산이 가장 큰 후보 뉴런을 은닉층에 1개의 층 개념으로 추가시키며, 망의 출력 값과 주어진 목표 출력 값 간의 오차를 최소화 시키도록 출력 뉴런에 연결된 가중치를 훈련시키는 방법이다. 이 방법을 원하는 해를 얻을 때까지 반복하여 망을 구성한다.

1-입력, 1-출력으로 구성된 문제를 풀기 위한 CC 알고리즘의 훈련 과정은 그림 1에 나타내었으며, 알고리즘은 다음과 같다.

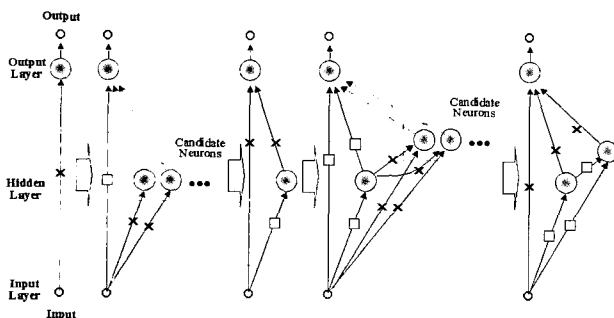


그림 1. CC 알고리즘(□ : 가중치 고정, × : 가중치 조절)  
Fig. 1. CC Algorithm(□ : Fixed Weight, × : Adjusted Weight)

#### Step 1. 초기화

은닉층이 없이 입력과 출력층으로만 구성되며, SSE (Sum Squared Error)를 최소화시키도록 입력-출력층 간 연결 가중치와 출력층 뉴런의 바이어스 가중치를 훈련시킨다.

#### Step 2. 후보 뉴런 훈련

입력층 및 기존의 은닉층 뉴런에서 후보 뉴런(일반적으로 8개 사용)으로 연결되지만, 후보 뉴런의 출력이 출력층

에는 연결되지 않는다, (“유령 또는 가상” 연결 상태). 이는 후보 뉴런의 작동이 신경망의 출력에 영향을 미치지 않음을 의미하며, 단지 후보 뉴런의 입력 연결 가중치를 훈련시키기 위한 공분산을 계산하기 위해 신경망의 출력 오차를 이용함을 의미한다. 후보 뉴런의 출력과 출력층 뉴런의 오차간의 공분산이 최대가 되도록 후보 뉴런의 입력 연결 가중치를 훈련(후보 뉴런의 입력 가중치 조절)시킨다.

#### Step 3. 공분산 최대 후보 뉴런을 은닉층에 추가

후보 뉴런 중에서 가장 최적의 뉴런(공분산이 최대가 되는 뉴런)을 선택하여 기존의 신경망에 추가하며, 은닉층에 추가된 후보 뉴런은 입력층과, 이전에 추가된 모든 은닉 뉴런과 연결되고, 이 연결의 가중치는 후보 뉴런 훈련에서 결정된 값을 가지며, 고정된다. 추가된 후보 뉴런과 출력층이 연결된다.

#### Step 4. 출력층 뉴런 가중치 훈련

신경망의 오차 (SSE)가 최소가 되도록 출력층에 연결된 모든 가중치(입력층, 기존에 추가된 은닉 뉴런들과 새로 추가된 은닉 뉴런으로부터)를 다시 훈련시킨다.

#### Step 5. 망의 성능 평가

훈련 결과 망의 성능을 평가하여 원하는 수준의 결과를 얻으면 훈련을 종료하고, 그렇지 않으면 Step 2로 복귀한다.

### 2.2 관련 연구

Fahlman과 Lebiere[4]의 CC 알고리즘 (CasCor)은 후보 뉴런 훈련에 “공분산 최대화”와 출력층 뉴런 훈련에 “오차 최소화” 기준을 적용하여 2단계 훈련을 수행하는 형태이다. 이에 반해 Prechelt[5]는 Fahlman과 Lebiere[4]의 “공분산 최대화” 기준을 적용하는 CC 알고리즘에 문제가 있음을 지적하고 후보 뉴런을 훈련시킬 때 “공분산 최대화” 대신 “오차 최소화” 기준을 적용하여 “오차 최소화”와 “오차 최소화”의 2단계 훈련을 수행하는 여러 가지 변형된 CC 알고리즘 (CasCor Family) 모델을 제안하였다. 또한 Lehtokangas[6]는 CC 알고리즘의 CC 구조를 유지하면서, 새로 추가되는 후보 뉴런에 연결된 입력-후보 뉴런, 후보 뉴런-출력층 연결 가중치를 “오차 최소화” 기준을 적용 역전파 알고리즘으로 1회 훈련시키는 모델을 제안하였다.

위에서 살펴본 바와 같이 CC 알고리즘에 대한 여러 가지 변형된 형태 또는 훈련 방법이 제안되었고, FFN에 대한 구성 알고리즘이 제안되었지만 패턴분류나 함수근사 문제에 있어서 CC 알고리즘보다 명확한 성능 향상을 나타낸 모델은 제시하지 못하였다[5-6].

따라서, 본 논문은 구성 알고리즘에서 가장 일반적으로 적용되고 있는 CC 알고리즘에 대해 모델의 구조와 연결 가중치 훈련 방법에 대한 변형인 CasCor 패밀리를 제시하고 모델의 성능을 평가하고자 한다.

## 3. CasCor 패밀리

Fahlman과 Lebiere[4]의 CC 알고리즘을 CasCor-all (CasCor 구조 및 출력층에 연결된 모든 가중치 재 훈련 알고리즘)로 칭하고 이를 변형시켜 제안하는 모델은 다음과 같다.

#### (1) CasCor-new

CasCor 구조이면서 출력층에 기존에 연결된 가중치는 고정시킨 상태에서 새로 은닉층에 추가되는 후보 뉴런과 출

력층 간에 연결된 가중치만 훈련시키는 알고리즘

(2) CasCorio-all

CoaCor에서 입력-출력층간 연결강도 제거, 출력층에 연결된 모든 가중치를 훈련시키는 알고리즘

(3) CasCorio-new

CasCorio 구조이면서 출력층에 기존에 연결된 가중치는 고정시킨 상태에서 새로 은닉층에 추가되는 후보 뉴런과 출력층 간에 연결된 가중치만 훈련시키는 알고리즘

4개 모델에 대해 2-입력, 1-출력을 가진 문제의 경우 은닉층에 4번째 뉴런이 추가된 시점에서의 신경망 모델의 구조는 그림 2와 같다.

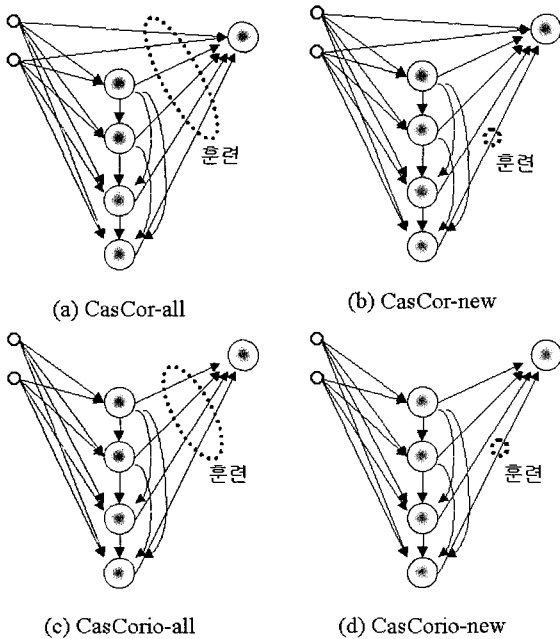


그림 2. CasCor 패밀리 신경망 구조

Fig. 2. Neural Network Architecture for CasCor Family

4개의 CasCor 패밀리 중에서 CasCor-all은 Fahlman과 Lebeire[4]가 제안한 CasCor 알고리즘이며, 나머지 3개는 새로 제안된 알고리즘으로 CosCor의 변형된 형태를 취함을 알 수 있다.

4. 실험 및 결과 분석

실험에 사용된 프로그램은 이미 적합성을 검증 받고 사용되고 있는 카네기 멜론 대학의 Crowder에 의해 개발된 프로그램[8]을 기본으로 하여 3장에서 제안된 모델들에 적합하도록 약간의 수정을 통해 생성되었다. 그러나 이 프로그램은 방대한 분량으로 그림 3에 일부분만을 제시하였다.

본 연구에는 그림4와 같이 패턴분류 문제에서는 전향망(Feedforward network : FFN)의 역전파(Backpropagation) 알고리즘으로 풀기에는 가장 어려운 2개 나선형(Two-spirals) 분류 문제[4], 함수근사 문제에서는 다양한 모델의 성능을 평가하는데 많이 이용되는 Mackey-Glass 미분함수[7]를 사용하였다.

```

.....*/
/* C implementation of the Cascade-Correlation learning algorithm. */
/*
/*
/* Written by:      R. Scott Crowder, III
/*                 School of Computer Science
/*                 Carnegie Mellon University
/*                 Pittsburgh, PA 15213-3890
/*
/*
/* Phone: (412) 258-8138
/*          internet: rsc@cs.cmu.edu
/*
/*
/* This code has been placed in the public domain by the author. As a matter of simple courtesy,
/* anyone using or adapting this code is expected to acknowledge the source. The author would like
/* to hear about any attempts to use this system, successful or not.
/* This code is a port to C from the original Common Lisp implementation written by Scott E.
/* Fahlman. (Version dated June 1 1990)
/*
/* For an explanation of this algorithm and some results, see "The Cascade-Correlation Learning
/* Architecture" by Scott E. Fahlman and Christian Lebiere in D. S. Touretzky (ed.), "Advances in
/* Neural Information Processing Systems 2", Morgan Kaufmann, 1990. A somewhat longer version
/* is available as CMU Computer Science Tech Report CMU-CS-90-103. Instructions for Fixing
/* this report are given at the end of this file.
/*
/* An example of the network set up file is provided at the bottom of this file.
/*
/* This code has been successfully compiled on the following machines.
/* DEC Station 3100 using the MIPS compiler version 1.31
/*
/* Sum 4 using the gcc compiler version 1.23
.....*/

```

그림 3. 실험에 적용된 프로그램 예.

Fig. 3. Example of simulation program

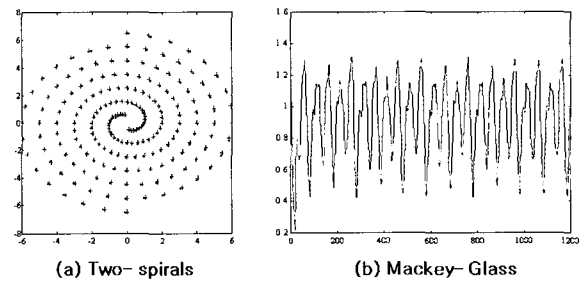


그림 4. 실험에 사용된 데이터

Fig. 4. Simulated Data

2개 나선형 데이터 생성 프로그램[9]을 이용하여 -0.5와 0.5로 입력 패턴을 분류하는 문제로 시험 데이터로 192개, 훈련 데이터로 192개를 생성하였다. 192개 훈련 데이터에 대해 모델을 훈련시키고, 시험 데이터 192개에 대해 분류 성능을 실험하였다. Mackey-Glass 미분함수 데이터[7]는  $\frac{dx(t)}{dt} =$

$$\frac{ax(t-\tau)}{1+x^{10}(t-\tau)} - bx(t) \text{ 이며, } a=0.2, b=0.1 \text{ 이고 } \tau=17 \text{ 이다.}$$

118 ≤ t ≤ 1117 범위에서, 입력 데이터는 {x(t-16), x(t-12), x(t-6), x(t)}에서의 x(t)로, 출력 데이터는 x(t+6)에서의 x(t)로 1000개의 데이터가 설정되었으며, 모델의 일반화 성능을 평가하기 위해 처음 500개를 훈련 데이터로 나머지 500개의 데이터가 시험 데이터로 준비되었다.

각각의 문제에 대해 3장에서 제시된 4개의 모델에 대해 은닉층-출력층 뉴런의 작동함수가 시그모이드-시그모이드인 경우와 시그모이드-선형인 경우로 구분하여 30회의 시행을 평균하여 비교 분석하였으며, 추가되는 최대 은닉 뉴런 수는 100개로 설정하였다.

2개의 나선형을 분류하는 문제에 대한 모델의 성능을 분석한 결과는 표 1에 제시되어 있다. 표 1의 결과로부터 기존의 Fahlman과 Lebiere[4]가 제안한 CasCor-all의 시그모이드-시그모이드 작동함수에 비해 나머지 7개 형태 모델들이 어떠한 결과를 나타내고 있는지는 모델의 단순성, 훈련시간과 모델의 성능으로 구분하여 평가하였다. 단순성은 은닉 뉴런수의 개수로, 훈련시간은 Epoch의 값으로, 모델 성능은 훈련 데이터와 시험 데이터의 실패율 또는 오차로 평가하였다. 이에 대한 결과는 표 2에 제시하였다.

표 1. 2개 나선형 분류 문제의 모델 성능  
Table 1. Model Performance of 2 Spiral Classification Problem

모델	작동함수		epoch	은닉 뉴런수	훈련 데이터 실패율	시험 데이터		
	은닉층	출력층				분류 실패 비트 수	실패율 (%)	
CasCor	all	시그모이드	시그모이드	1925	17	0%	11.6	6.0417
		시그모이드	선형	2306	25	0%	12.6	6.5625
	new	시그모이드	시그모이드	1653	14	0%	8.8	4.5833
		시그모이드	선형	2285	24	0%	12.2	6.3542
CasCorio	all	시그모이드	시그모이드	1969	19	0%	12.7	6.6146
		시그모이드	선형	2635	29	0%	13.9	7.2396
	new	시그모이드	시그모이드	1899	31	0%	12.7	6.6146
		시그모이드	선형					실패

표 2. 패턴분류의 성능 평가  
Table 2. Performance Evaluation for Pattern Classification

모델	작동함수		성능 평가 (순위)				합계/ 순위	
	은닉층	출력층	훈련 시간 (epoch)	모델 단순성 (은닉 뉴런수)	훈련 데이터 실패율	시험 데이터 실패율		
CasCor	all	시그모이드	시그모이드	3	2	1	2	8/2
		시그모이드	선형	6	5	1	4	16/6
	new	시그모이드	시그모이드	1	1	1	1	4/1
		시그모이드	선형	5	4	1	3	13/3
CasCorio	all	시그모이드	시그모이드	4	3	1	6	14/4
		시그모이드	선형	7	6	1	7	21/7
	new	시그모이드	시그모이드	2	7	1	5	15/5
		시그모이드	선형					실패

패턴분류 문제에 있어서, 다음과 같은 결론을 얻을 수 있었다.

- (1) 훈련 데이터에 대한 성능은 4개 모델 모두 작동함수를 시그모이드-시그모이드 형태로 사용하는 경우와 시그모이드-선형으로 사용하는 경우 동일하게 100% 분류 성공률을 나타내었다. (단, CasCorio-new의 시그모이드-선형 작동함수 모델 제외)
- (2) 모델의 일반화 성능 (시험 데이터로 평가)에 대해서는 CasCor 모델에서 출력층에 연결된 모든 연결 가중치를 다시 훈련시키는 CasCor-all 보다 새로 추가되는 뉴런과 출력층간 연결 가중치만 훈련시키는 CasCor-new가 보다 좋은 결과를 나타내었다. 이는 기존에 추가된 은닉 뉴런과 출력층간 연결 강도는 이미 훈련이 수행되어 신경망 출력의 오차를 감소시키기 위해 각각의 역할을 수행하고 있으며, 이를 다시 훈련시킴으로서 Fahlman과 Lebiere[4]가 지적했듯이 FFN의 구성 알고리즘인 DNC(Dynamic Node Creation) 알고리즘[10]에서 빈번히 발생하는 “움직이는 목표 문제” (Moving Target Problem)로 원하는 목표를 찾지 못하고 결국 나쁜 결과를 나타내는 것으로 판단된다.
- (3) 2개 나선형 분류 문제에 있어서 CasCorio-new의 경우 은닉 뉴런을 100개 추가하고도 훈련 데이터의 오차 기준 (훈련 데이터의 분류 실패 비트 수 = 0)을 만족하지 못하여 훈련에 실패하였다.

Mackey-Glass 미분함수에 대한 모델의 일반화 성능은 표 3에 기술되어 있다. 함수근사 문제에 있어서, 제안된 모델

이 기존의 CasCor 모델보다 좋은 결과를 얻고 있는지 모델의 단순성, 훈련시간과 모델의 성능으로 구분하여 평가하였으며 표 4에 제시되어 있다.

표 3. Mackey-Glass 미분함수의 일반화 성능  
Table 3. Generalization Performance of Mackey-Glass Differential Function

모델	작동함수		epoch	은닉 뉴런수	훈련 데이터 오차	시험 데이터 오차	
	은닉층	출력층					
CasCor	all	시그모이드	시그모이드	4	0	121.5845	121.0940
		시그모이드	선형	1066	7	0.2552	0.2517
	new	시그모이드	시그모이드	5	0	118.3449	117.8773
		시그모이드	선형	943	6	0.2542	0.2507
CasCorio	all	시그모이드	시그모이드	46	1	123.9372	123.4418
		시그모이드	선형	1007	12	0.2522	0.2492
	new	시그모이드	시그모이드	1053	27	172.5087	171.3631
		시그모이드	선형				실패

표 4. 함수근사의 성능 평가  
Table 4. Performance Evaluation for Function Approximation

모델	작동함수		성능 평가 (순위)				합계/ 순위	
	은닉층	출력층	훈련 시간 (epoch)	모델 단순성 (은닉 뉴런수)	훈련 데이터 오차	시험 데이터 오차		
CasCor	all	시그모이드	시그모이드	1	1	5	5	12/2
		시그모이드	선형	5	7	3	3	18/5
	new	시그모이드	시그모이드	1	2	1	1	11/1
		시그모이드	선형	4	4	2	2	12/2
CasCorio	all	시그모이드	시그모이드	3	3	6	6	18/5
		시그모이드	선형	6	5	1	1	13/4
	new	시그모이드	시그모이드	7	6	7	7	27/7
		시그모이드	선형					실패

함수근사 문제에 있어서 다음과 같은 결론을 얻을 수 있었다.

- (1) 작동함수를 시그모이드-시그모이드 형태 보다는 시그모이드-선형을 사용하는 경우 월등한 성능 향상을 나타내었다. 이는 Fahlman과 Lebiere[4]의 CasCor 알고리즘에서 이용한 시그모이드-시그모이드 작동함수를 사용해 공분산 최대화 기준을 적용시 발생하는 문제로 나타난 결과이다. 왜냐하면, 두 변량의 공분산 또는 상관계수는 두 변량의 일차 중속 또는 선형 중속에 대한 측도로 사용되며, 두 변량 사이에 강한 연관관계가 있음에도 불구하고 그 관계가 선형이 아니면 극단적인 경우 상관계수가 0이 되는 경우도 발생한다. 그러므로, 두 변량 (출력층 뉴런과 후보 뉴런) 모두 비선형인 시그모이드 작동함수를 사용하면 공분산이 왜곡되는 경향이 발생할 수 있다. 따라서 2개 변수 중 출력층 뉴런의 작동함수를 선형으로 사용하는 것이 이 문제를 해결할 수 있는 대안이 될 수 있다.
- (2) 시그모이드-선형 작동함수 사용시, CasCor 모델 출력층 뉴런의 연결 가중치 모두를 훈련시키는 CasCor-all 보다 새로 추가된 은닉뉴런에 관련된 연결 가중치만 훈련시키는 CasCor-new가 보다 나은 결과를 얻을 수 있다.
- (3) 시그모이드-선형 작동함수 사용시, 입력-출력층 사이에

연결된 CasCor 모델보다 입력·출력층 사이에 연결이 안 된 CasCorio 모델이 보다 좋은 결과를 나타내었다. 따라서 CasCorio 모델이 CasCor 모델보다 단순해지고, 훈련 시간 단축 및 보다 좋은 결과를 나타내는 모델로 판단할 수 있다.

- (4) 2개 나선형 분류 문제와 동일하게 함수근사 문제에 있어서도 CasCorio-new의 경우 은닉 뉴런을 100개 추가하고도 훈련 데이터의 오차 기준을 만족하지 못하여 훈련에 실패하였다.

### 5. 결론 및 향후과제

본 논문은 구성 알고리즘에서 가장 일반적으로 사용되고 있는 Fahlman과 Lebiere[4]가 제안한 공분산 최대화, 오차 최소화 CC 알고리즘 (CasCor)의 구조 및 훈련방법을 변경한 다양한 모델을 제시하였다.

연구 결과, 패턴분류 문제에 대해서는 시그모이드-시그모이드 작동함수를 사용하고, 새로운 은닉 뉴런과 연결된 출력층의 가중치만 훈련시키는 CasCor-new 모델이 가장 적합한 모델로 판단된다. 또한 함수근사 문제에 대해서는 모델의 구조에 상관없이 시그모이드-선형 작동함수를 사용하는 모델이 보다 월등한 결과를 나타내었으며, 그 중에서도 CasCorio 모델이 가장 적합한 모델로 판단된다.

CasCor 알고리즘의 구조 측면에서 볼 때, 보다 단순한 형태를 취하도록 모델을 변경할 수도 있을 것으로 판단된다. 따라서, 추후 이 분야에 대한 연구를 수행할 것이다.

### 참 고 문 헌

[1] T. Y. Kwok and D. Y. Yeung, "Constructive Algorithms for Structure Learning in Feedforward Neural Networks for Regression Problems," IEEE Trans. on Neural Networks, Vol. 8, No. 3, pp. 630-645, 1997.

[2] J. Ghosh and K. Tumer, "Structural Adaptation and Generalization in Supervised Feed-forward Networks," Journal of Artificial Neural Networks, Vol. 1, No. 4, pp. 431 - 458, 1994.

[3] J. Moody, "Prediction Risk and Architecture Selection for Neural Networks," Theory and Pattern Recognition Applications, NATO ASI Series, F, pp. 147-165, Springer-Verlag, 1994.

[4] S. E. Fahlman and C. Lebiere, "The Cascade Correlation Learning Architecture," Advances in Neural Information Processing Systems II, pp. 525-532, 1990.

[5] L. Prechelt, "Investigation of the CasCor Family of Learning Algorithms," Neural Networks, Vol. 10, No. 5, pp. 885 - 896, 1997.

[6] M. Lehtokangas, "Modeling with Constructive Backpropagation," Neural Networks, Vol. 12, pp. 707-716, 1999.

[7] E. Littmann and H. Ritter, "Cascade Network Architectures," Proc. Intern. Joint Conference on Neural Networks, Vol. II, pp. 398-404, 1992.

[8] R. S. Crowder, "CASCOR: Lisp and C Implementations of CasCade Correlation," ftp://ftp.cs.cmu.edu/afs/cs.cmu.edu/project/connect/code/supported/

[9] Cascor1, "http://www.cs.cmu.edu/afs/cs/project/ai-repository/ai/areas/neural/systems/cascor/c/cascor1.c

[10] T. Ash, "Dynamic Node Creation in Backpropagation Neural Networks," Connection Science, Vol. 1, No. 4, pp. 365-375, 1989.

### 저 자 소 개



#### 최명복(Myong-Bok Choi)

1992년 : 호서대학교 전자계산학과(학사)  
 1994년 : 아주대학교 컴퓨터공학과(석사)  
 2001년 : 아주대학교 컴퓨터공학과(박사)  
 1997~현재 원주대학 행정전산과(부교수)  
 2003. 9~현재 한국정보처리학회 전문대학전산교육연구회 운영위원  
 2003. 9~현재 정보과학회 전문대학 학회

지·논문지 편집위원

2004. 1~현재 한국컴퓨터산업교육학회 학회지 편집위원

2004. 1~현재 한국인터넷방송/TV학회 회원관리이사

관심분야 : 지능형 정보검색, 퍼지응용, 지식표현, 신경망, 지능형 교통제어, 소프트웨어 공학



#### 이상운(Sang-Un Lee)

1987년 : 한국항공대학교 항공전자공학(공학사)

1997년 : 경상대학교 컴퓨터과학과(석사)

2001년 : 경상대학교 컴퓨터과학과(박사)

1992~2003. 2 : 국방품질관리소 항공전자 장비 및 소프트웨어 품질보증 담당

2003. 3~2004. 2 : 도립 강원전문대학 컴

퓨터응용과 전임강사

2004. 3~현재 국립 원주대학 여성교양과 전임강사

관심분야 : 소프트웨어공학 (Methodology, Metrics, Testing, Reliability, Project Management, Architecture, CBD), 신경망, 뉴로-퍼지