'96 대한전기학회 하계학술대회 논문집 (1996. 7. 22~24)

Langevine 경쟁학습 신경회로망의 확산성과 대역 최적화 성질의 근사 해석

석 진 욱° 조 성 원 최 경 삼
홍익 대학교 공과대학 전자·전기 공학부

An Informal Analysis of Diffusion, Global Optimization Properties in Langevine Competitive Learning Neural Network

Jinwuk Seok Seongwon Cho Gyung-Sam Choi
School. of Electronic and Electric Eng., Hong Ik University

Abstract

In this paper, we discuss an informal analysis of diffusion, global optimization properties of Langevine competitive learning neural network. In the view of the stochastic process, it is important that competitive learning gurantee an optimal solution for pattern recognition. We show that the binary reinforcement function in Langevine competitive learning is a brownian motion as Gaussian process, and construct the Fokker-Plank equation for the proposed neural network. Finally, we show that the informal analysis of the proposed algorithm has a possiblity of globally optimal solution with the proper initial condition.

1. 서 론

본 연구에서는 기존의 경쟁학습 신경회로망보다 실험적으로 학습성능이 우수한 것으로 알려지고 하드웨어 구현에 편리한 Langevine 경쟁학습 신경회로망의 확산성과 대역 최소화 성질을 근사적으로 분석해 본다[1]. 본 논문에서는 먼저 Langevine 경쟁학습 신경회로망의. 학습 알고리즘의 Fokker-Plank 방정식을 유도하고 이 방정식을 이용하여 Langevine 경쟁학습 신경회로망 학습 방정식의 Fokker-Plank 방정식이 $L[0,1]$공간 상에 Contraction되어 있음을 보인다. 마지막으로 본 논문에서 사용된 신경회로망 알고리즘에 의해 구축된 퍼지 전문가 시스템을 전력계통 지락사고 패턴에 적용하여, 실험적으로 Langevine 경쟁학습 신경회로망에 의한 퍼지 전문가 시스템의 구축방법이 보다 좋은 성능을 나타냄을 보인다.

2. Langevine 경쟁학습 신경회로망

Langevine 경쟁학습 신경회로망의 알고리즘은 다음과 같다[1].

1. 뉴런 r이 winner 이면 :

$$w^r_{t+1} = w^r_t + \varepsilon(v_t - w^r_t) + \eta(\lambda, \delta, t, v_t, w^r_t)$$

2. 뉴런 r이 winner가 아니면 :

$$w^r_{t+1} = w^r_t$$

여기에서 w^r_t는 시간 t일 때 r번째 Weight vector를 의미하며 ε은 시간에 대하여 불변인 일정적용 이득, v_t는 시간 t일 때 입력 Vector, 그리고 $\eta(\lambda, \delta, t, v_t, w^r_t)$는 이진 강화함수 항으로서 다음과 같이 구성된다.

$$\eta(\lambda, \delta, t, v_t, w^r_t) = \lambda\delta(t)sgn(v_t - w^r_t) \tag{1}$$

식 (1)에서 λ는 이진 강화함수의 이득이며, $\delta(t)$는 1혹은 0을 반환하는 Random Variable로서 본 논문에서는 시간에 종속인 $P_\delta(t)$의 확률로 1을 반환하며, $sgn(v_t - w^r_t)$은 다음을 만족하는 함수이다.

$$sgn(v_t - w^r_t) = \begin{cases} 1 & v_t - w^r_t \geq 0 \\ 0 & v_t - w^r_t < 0 \end{cases} \tag{2}$$

이때 제안한 경쟁학습 신경회로망의 성능지표 J_t는 다음과 같다.

$$J_t = \frac{1}{2}\int_{B_t}\sum_r^W \| v_t - w^r_t \|^2 dP(v_t) \tag{3}$$

한편 시간 t에서의 성능지표 J_t에 대한 Directional Derivation h_t를 J_t의 Gradient로 놓으면

$$h_t = \nabla J_t = -(v_t - w_t^r)$$

따라서 Directional Derivation h_t의 Descent Direction h^D_t를 Theorem 1과 같이 놓으면 h^D_t는 J_t를 최소화 하는 방향으로 w^r_t를 보낸다.

Theorem 1 : Descent Direction h^D_t가 다음과 같이 정의되면

$$h^D_t = \varepsilon h_t + \lambda\delta(t)sgn(v_t - w^r_t) \qquad \varepsilon > 0$$

학습 방정식 $w^r_{t+1} = w^r_t + h^D_t$는 J_t를 최소화 하는 방향으로 w^r_t를 보낸다.

proof :

이진 강화 함수와 J_t의 Weight vector w^r_t에 대한 Gradient ∇J_t와 h^D_t의 스칼라 적이 0보다 작다는 것과 위 명제는 동치이다. 따라서 이진 강화함수와 ∇J_t의 스칼라 적을 살펴보면 이진 강화함수의 정의에 의해 이진 강화함수 n차원 각 Component들은 크기가 1 혹은 0이며 부호가 ∇J_t의 대응되는 Component와 반대이므로

$$-n \leq \langle \nabla J_t, \lambda\delta(t)sgn(v_t - w^r_t)\rangle \leq 0$$

그러므로

$$\langle \nabla J_t, \varepsilon h_t + \lambda\delta(t)sgn(v_t - w^r_t)\rangle$$
$$= -\varepsilon\|\nabla J_t\|^2 + \langle \nabla J_t, \lambda\delta(t)sgn(v_t - w^r_t)\rangle < 0 \qquad \blacksquare$$

3. Langevine 경쟁학습 신경회로망의 Fokker-Plank 방정식 유도

Langevine 경쟁학습 신경회로망 이진 강화함수 항은 Gaussian Process이므로 Langevine 경쟁학습 신경회로망의 학습 방정식은 다음과 같이 일반적인 확률 미분 방정식(SDE)의 형태를 가지게 된다[6].

$$\partial w^r_t = -\varepsilon\nabla J(w^r_t)dt + \eta(t, \cdot)dt \tag{4}$$

이진 강화함수 항 $\eta(t, \cdot)$ 는 "Gaussian Process"이므로 다음의 성질을 만족한다[6].

$$E(\eta(t,\cdot))=\langle\eta(t,\cdot)\rangle=0$$
$$E(\eta(t,\cdot)\eta(t,\cdot))=\langle\eta(t,\cdot)\eta(s,\cdot)\rangle=\lambda^2 p_\delta s\,\delta(t-s)$$
$$E(\prod_{i=1}^{n}\eta(t_i,\cdot))=\langle\prod_{i=1}^{n}\eta(t_i,\cdot)\rangle=\begin{cases}0 & : l \text{ is odd}\\ \sum\prod\lambda^2 p_\delta\delta(t-s) & : l \text{ is even}\end{cases}$$

시간 t에서 r번째 Weight vector w^r_t의 Transition에 대한 확률밀도 함수를 $\rho(w^r_t,t)$라 놓고 $\rho(w^r_t,t)$의 시간에 대한 변화율을 고려하면,

$$\frac{\partial\rho(w^r_t,t)}{\partial t}=-\frac{\partial}{\partial w^r_t}\left[\frac{\partial w^r_t}{\partial t}\rho(w^r_t,t)\right] \tag{5}$$

식 (1)을 식 (2)에 대입하면

$$\frac{\partial\rho(w^r_t,t)}{\partial t}=-\frac{\partial}{\partial w^r_t}[(-\nabla J(w^r_t)+\eta(t,\cdot))]\rho(w^r_t,t) \tag{6}$$

여기서 변위에 대한 Foward 선형 미분 연산자 L_f를 다음과 같이 정의한다.

Definition 1 : 변위에 대한 Foward 선형 미분 연산자 L_f는 다음과 같이 정의된다.

$$L_f\rho(x,t)\equiv\frac{\partial}{\partial x}[A(x)\rho(x,t)]=\left[\frac{\partial A(x)}{\partial x}+A(x)\frac{\partial}{\partial x}\right]\rho(x,t)$$

Definition 1에 의해 정의된 연산자 L_f를 식 (6)에 도입하면

$$L_f\rho(w^r_t,t)=\varepsilon\frac{\partial}{\partial w^r_t}[\nabla J(w^r_t)\rho(w^r_t,t)] \tag{7}$$
$$=\varepsilon\left[\frac{\partial\nabla J(w^r_t)}{\partial w^r_t}+\nabla J(w^r_t)\frac{\partial}{\partial w^r_t}\right]\rho(w^r_t,t)$$

식 (7)을 식 (6)에 대입하면

$$\frac{\partial\rho(w^r_t,t)}{\partial t}=L_f\rho(w^r_{t,t})-\frac{\partial\rho(w^r_t,t)}{\partial w^r_t}\eta(t,\cdot) \tag{8}$$

편미분 방정식 (8)을 시간에 대하여 풀기 위해 식 (8)의 해의 한 후보를 다음과 같이 놓는다.

$$\rho(w^r_t,t)=e^{tL_f}\xi(w^r_t,t) \tag{9}$$

식 (9)에서 $\xi(w^r_t,t)$는 시간 t에 종속인 함수이므로 w^r_t에 대한 미분 값은 0이다. 식 (9)을 t에 대하여 편미분하면

$$\frac{\partial\xi(w^r_t,t)}{\partial t}=-L_f(w^r_t)e^{-tL_f}\rho(w^r_t,t)+e^{-tL_f}\frac{\partial\rho(w^r_t,t)}{\partial t}$$
$$=e^{-tL_f}[-L_f(w^r_t)\rho(w^r_t,t)+L_f(w^r_t)\rho(w^r_t,t) \tag{10}$$
$$-\eta(t,\cdot)\frac{\partial\rho(w^r_t,t)}{\partial w^r_t}]$$
$$=-\eta(t,\cdot)e^{-tL_f}\left[\frac{\partial e^{tL_f}}{\partial w^r_t}\xi(w^r_t,t)+\frac{\partial\xi(w^r_t,t)}{\partial w^r_t}e^{tL_f}\right]$$

식 (10)에서 $\xi(w^r_t,t)$의 t에대한 미분연산자 $K(w^r_t,t)$를 다음과 같이 놓는다.

$$K(w^r_t,t)=-\eta(t,\cdot)e^{-tL_f(w^r_t)}g(w^r_t,t)\frac{\partial}{\partial w^r_t}e^{tL_f(w^r_t)}$$

$\xi(w^r_t,t)$를 적분형 Taylor급수로 놓으면

$$\xi(w^r_t,t+\Delta t)=\xi(w^r_t,t)+\int_t^{t+\Delta t}dt_1K(w^r_t,t)\xi(w^r_t,t)$$
$$=[1+\int_t^{t+\Delta t}dt_1K_1(w^r_t,t) \tag{11}$$
$$+\frac{1}{2!}\int_t^{t+\Delta t}dt_1\int_t^{t+\Delta t}dt_2K_1(w^r_t,t)K_2(w^r_t,t)$$
$$+o(w^r_t,t)]\xi(w^r_t,t)$$

$\xi(w^r_t,t+\Delta t)$의 평균을 식 (11)에서 구하게 되면

$$E\xi(w^r_t,t+\Delta t)=[1+\int_t^{t+\Delta t}dt_1EK_1(w^r_t,t)$$
$$+\frac{1}{2!}\int_t^{t+\Delta t}dt_1\int_t^{t+\Delta t}dt_2E(K_1(w^r_t,t)K_2(w^r_t,t)) \tag{12}$$
$$+Eo(w^r_t,t)]E\xi(w^r_t,t)$$

식 (12)에서 $K(w^r_t,t)$의 2차 모멘트 까지의 평균을 구하면

$$E(K(w^r_t,t)\xi(w^r_t,t))$$
$$=E\eta(t,\cdot)e^{-tL_f(w^r_t)}\frac{\partial}{\partial w^r_t}e^{tL_f(w^r_t)}E\xi(w^r_t,t)=0$$
$$E(K(w^r_t,t)K(w^r_t,s)g(w^r_t,t))$$
$$=E(\eta(t,\cdot)\eta(s,\cdot))e^{-tL_f(w^r_t)}\frac{\partial}{\partial w^r_t}e^{tL_f(w^r_t)}$$
$$\cdot e^{-tL_f(w^r_t)}\frac{\partial}{\partial w^r_t}e^{tL_f(w^r_t)}E\xi(w^r_t,t)$$
$$=\lambda^2p_\delta\delta(t-s)e^{-2tL_f(w^r_t)}\frac{\partial^2}{\partial^2w^r_t}e^{tL_f(w^r_t)}E\xi(w^r_t,t)$$

위 결과를 통해 Talyor급수의 2차항에 대한 평균을 구하면

$$\frac{1}{2!}\int_t^{t+\Delta t}dt_1\int_t^{t+\Delta t}dt_2E(K_1(w^r_t,t)K_2(w^r_t,t)\xi(w^r_t,t))$$
$$=\frac{1}{2!}\int_t^{t+\Delta t}dt_1\int_t^{t+\Delta t}\lambda^2p_\delta\delta(t_1-t_2)e^{-t_1L_f(w^r_t)}$$
$$\cdot\frac{\partial}{\partial w^r_t}e^{(t_1-t_2)L_f(w^r_t)}\frac{\partial}{\partial w^r_t}e^{-t_2L_f(w^r_t)}E\xi(w^r_t,t)$$
$$=\frac{1}{2}\lambda^2p_\delta\Delta te^{-tL_f(w^r_t)}\frac{\partial^2}{\partial^2w^r_t}e^{tL_f(w^r_t)}E\xi(w^r_t,t)$$

한편 고차항에서는 앞에서 살펴본 바와 같이 차수 n이 홀수이면 0, 2n이면 $(\Delta t)^n$에 비례한다. 따라서 $\Delta t\mapsto 0$의 극한의 경우 2차항까지만 고려한다면

$$\lim_{\Delta t\to0}\frac{E\xi(w^r_t,t+\Delta t)-E\xi(w^r_t,t)}{\Delta t}=\frac{\partial E\xi(w^r_t,t)}{\partial t}$$
$$=\frac{1}{2}\lambda^2p_\delta e^{-tL_f(w^r_t)}\frac{\partial^2}{\partial^2w^r_t}e^{tL_f(w^r_t)}E\xi(w^r_t,t)$$

$P(w^r_t,t)$를 $\rho(w^r_t,t)$의 ensemble에서 취한 평균으로 놓으면 : $P(w^r_t,t)=E\rho(w^r_t,t)=e^{tL_f(w^r_t)}E\xi(w^r_t,t)$

따라서 다음의 Fokker-Plank방정식을 얻을 수 있다.

$$\frac{\partial P(w^r_t,t)}{\partial t}=L_fe^{tL_f}E\xi(w^r_t,t)+e^{tL_f}\frac{\partial E\xi(w^r_t,L_f)}{\partial t}$$
$$=L_fP(w^r_t,t)+\frac{1}{2}\lambda^2p_\delta\frac{\partial^2}{\partial^2w^r_t}e^{tL_f(w^r_t)}E\xi(w^r_t,t)$$
$$=\frac{\partial}{\partial w^r_t}[\varepsilon\nabla J(w^r_t)P(w^r_t,t)]+\frac{1}{2}\lambda^2p_\delta\frac{\partial^2}{\partial^2w^r_t}P(w^r_t,t)$$

4. Fokker-Plank 방정식을 사용한 Langevine 경쟁학습 신경회로망의 대역 최소화 근사 해석

확률 미분방정식을 통해 Transition 확률이 유일한 Transition 확률분포의 해 $\pi\propto\exp(-J(w^r_t x_t)/\sigma^2(t))$가 존재하면 π는 $J(w^r_t x_t)$의 대역 최소점으로 w^r_t을 보내게 된다 [7]. 따라서, 유도된 Fokker-Plank방정식이 어떤 조건에서 Contraction될 수 있는가를 증명하면 Langevine 경쟁학습의 학습 방정식이 최적해를 찾을 수 있는가와 최적해를 찾기 위한 조건을 얻을 수 있다.

Definition 2: Fokker-Plank방정식에 대한 Infinitestimal Operator $\mathcal{L}^o_t$를 다음과 같이 정의한다.

$$\mathcal{L}^o_t=-\frac{\partial}{\partial w^r_t}a(w^r_t,v_t)+\frac{1}{2}\sigma(t)\frac{\partial^2}{\partial^2w^r_t}$$

$a(w^r_t,v_t)=-\varepsilon\nabla J(w^r_t)$, $\sigma(t)=\lambda^2p_\delta(t)\,t>0$이며 Fokker-Plank 방정식은 $\frac{\partial P(w^r_t,t)}{\partial t}-\mathcal{L}^o_tP(w^r_t,t)=0$ 이다.

Definition 3: Infinitestimal Operator $\mathcal{L}^o_t$는 $P(\cdot)\in L[0,1]$인 임의의 함수 f에 대하여 다음과 같이 Quasi-Norm을 정의한다.

$$\mathcal{L}_t=\|\mathcal{L}^o_t\|:=\sup_{t_0\le\tau\le t}\mathcal{L}^o_tP(w^r_t,t)/P(w^r_t,t)$$

Infinitestimal Operator $\mathcal{L}^o_t$의 Quasi-Norm $\mathcal{L}_t$는 오직 다음의 Norm 성질만 만족한다.

$$\mathcal{L}_t=0\quad\forall P(w^r_t,t)\in L[0,1]\quad\sup_{t_0\le\tau\le t}\mathcal{L}^o_tP(w^r_t,t)=0$$

Theorem 2: Langevine 경쟁학습 신경회로망의 Fokker-Plank방정식 $\partial_t P(w^r{}_i, t) = \mathcal{L}_i^o P(w_i^r, t)$은 Infinitestimal Operator $\mathcal{L}_i^o$의 Quasi-Norm의 시간에 대한 적분이 다음을 만족하면 $L[0,1]$에서 Contraction되어 있다.

$$\int_{t_0}^{t} \mathcal{L}_s ds \le \log\left(1 - \frac{\varepsilon}{P(w^r{}_{t_0}, t_0)}\right) \qquad \forall \varepsilon \in (0, P(w^r{}_{t_0}, t_0))$$

Proof : Lanevine 경쟁학습 신경회로망의 Fokker- Plank방정식은 다음과 같다.

$$\frac{\partial P(w^r{}_i, t)}{\partial t} = \mathcal{L}_i^o P(w_i^r, t) \le \mathcal{L}_t P(w_i^r, t) \tag{13}$$

$\mathcal{L}_s$를 다음과 같이 정의하고

$\mathcal{L}_s = \sup_{s-\varepsilon < \kappa < s+\varepsilon} \mathcal{L}_s^o P(w_s^r, s) / P(w_s^r, s) \quad \forall P(w_s^r, s) \in L[0,1]$

식 (13)의 양변에 $\exp(-\int_{t_0}^{t} \mathcal{L}_s ds)$을 곱한다.

$$\frac{\partial P(w^r{}_i, t)}{\partial t} \exp\left(-\int_{t_0}^{t} \mathcal{L}_s ds\right) - \mathcal{L}_t P(w_i^r, t) \exp\left(-\int_{t_0}^{t} \mathcal{L}_s ds\right) \le 0$$

$$\frac{\partial}{\partial t} P(w^r{}_i, t) \exp\left(-\int_{t_0}^{t} \mathcal{L}_s ds\right) \le 0 \tag{14}$$

식 (14)를 Lebesgue 적분하면

$$\int_{P(w^r_{t_0}, t_0)}^{P(w^r{}_i, t)\exp(-\int_{t_0}^{t} \mathcal{L}_s ds)} \partial P(w^r{}_i, t) \exp\left(-\int_{t_0}^{t} \mathcal{L}_s ds\right) \le 0$$

$$P(w^r{}_i, t) \exp\left(-\int_{t_0}^{t} \mathcal{L}_s ds\right) \le P(w^r_{t_0}, t_0)$$

$$P(w^r{}_i, t) \le P(w^r_{t_0}, t_0) \exp\left(\int_{t_0}^{t} \mathcal{L}_s ds\right) \tag{15}$$

식 (15)의 양변에 초기 Transient 확률 $P(w^r_{t_0}, t_0)$를 빼고 Norm을 취하면 다음의 관계식이 성립한다.

$$P(w^r{}_i, t) - P(w^r_{t_0}, t_0) \le P(w^r_{t_0}, t_0) \exp\left(\int_{t_0}^{t} \mathcal{L}_s ds\right) - P(w^r_{t_0}, t_0)$$

$$\begin{aligned} \|P(w^r{}_i, t) - P(w^r_{t_0}, t_0)\| &\le \|P(w^r_{t_0}, t_0)[\exp(\int_{t_0}^{t} \mathcal{L}_s ds) - 1]\| \\ &\le \|P(w^r_{t_0}, t_0)\| \cdot \|1 - \exp(\int_{t_0}^{t} \mathcal{L}_s ds)\| \end{aligned}$$

가정에서 $\int_{t_0}^{t} \mathcal{L}_s ds \le \log\left(1 - \frac{\varepsilon}{P(w^r{}_{t_0}, t_0)}\right)$ 이므로

$$\begin{aligned} \|P(w^r{}_i, t) - P(w^r_{t_0}, t_0)\| &\le \|P(w^r_{t_0}, t_0)\| \cdot \|1 - \exp(\int_{t_0}^{t} \mathcal{L}_s ds)\| \\ &\le \|P(w^r_{t_0}, t_0)\| \cdot \|1 - (1 - \frac{\varepsilon}{P(w^r_{t_0}, t_0)})\| \\ &\le \varepsilon \end{aligned}$$

가정에서 ε은 $\varepsilon \in (0, P(w^r{}_{t_0}, t_0))$를 만족하는 임의의 양수이므로 Langevine 경쟁학습 신경회로망의 Fokker- Plank방정식 $\partial_t P(w^r{}_i, t) = \mathcal{L}_i^o P(w_i^r, t)$은 $L[0,1]$에서 Contraction되어 있다. ■

5. 실험 결과

5.1 퍼지 집합의 멤버쉽함수 유도

퍼지 집합들의 멤버쉽함수들을 계산하기 위해서 우선 경쟁학습 알고리즘을 사용하여 menbership function들의 center들을 결정하고, 멤버쉽함수들의 width는 각 멤버쉽함수들이 서로 어느정도 (보통 25 %) 겹치도록 택한다. 멤버쉽함수들의 center를 구하기 위해, 조건부나 결론부에 사용되는 각각의 퍼지 변수에 대해서 경쟁학습을 사용하여 학습데이타를 여러개의 cluster로 나누어서 cluster의 mean값을 m멤버쉽함수들의 center로 정한다.

5.2 입출력 데이타

전송선로의 고장진단을 하기 위해서 사용되는 입력데이타는 전자 과도 현상 해석 프로그램 EMTP 로부터 얻어진 전류의 순시치를 전처리하여 생성한 A상 B상 C상 O상의 정규화된 실효치전류를 사용한다. 성능의 적절한 평가를 위해 임의의 고장점, 고장저항, 고장발생각을 가진 전력계통을 모델화하여 고장데이타를 얻었다. 출력패턴은 먼저 정상상태와 A상지락사고, AB상지락사고, AB선간단락사고의 4가지 경우에 대한 고장진단을 수행하는 SYSTEM을 구성하였다.

5.3. 실험결과

실험은 가장 단순한 경쟁학습 알고리즘인 SCL을 이용한 방법과, 본 논문에서 사용한 방법,그리고 신경회로망을 이용하지 않는 방법에 대해서 수행되었다. 학습과 테스트에는 각 고장의 상태별로 38개씩 총 152개의 데이타를 사용하였다. 실험은 먼저 학습용 데이타로 퍼지 집합을 생성하고 생성된 집합을 이용하여 rule을 추출하는 순서로 진행하였다. 단순경쟁학습에는 학습률은 $C(t) = 3140/(t+5500)$을 이용하였으며 그리고 신경회로망에서는 모두 각 100 epoch씩 학습을 시켰다. 또 신경회로망을 이용하지 않는 방법에서는 CDM의 임계치를 0.8로 설정하였다. 각각의 방법에 의해 구축된 지식기반을 사용한 퍼지시스템의 분류정확도 측정 실험 결과는 아래의 <표1>과 같다.

지식추출방법	정확도 (%)	
	학습(train)	테스트(test)
단순경쟁학습	86.8	84.8
Langevine 경쟁학습 신경회로망	88.2	86.2
신경회로망을 이용하지 않는 방법	87.5	85.5

[표 1]

7. 결 론

본 논문에서는 Langevine 경쟁학습 신경회로망 학습방정식의 Fokker-Plank방정식을 유도하고 유도된 Fokker-Plank방정식이 Contraction되어 있음을 보여 Langevine 경쟁학습 신경회로망이 대역 최소점에 확률적으로 접근할 수 있음을 보였다. 또한 본 논문에서 사용한 경쟁학습 신경회로망을 퍼지 전문가 시스템 자동구축에 응용하여 기존의 경쟁학습 신경회로망과 신경회로망을 이용하지 않은 퍼지 전문가 시스템과 고장 패턴에 대한 패턴 분류 실험에서 본 논문에서 제안한 방법이 보다 높은 성능을 나타내었다. 이는 일정적응 이득과 이진 강화함수를 가진 경쟁학습 신경회로망의 학습 방정식에 의해 생성된 수열의 극점 $\widehat{w^r}_i$이 단순 경쟁학습보다 더 낮은 J_i에 도달하여 전체적인 퍼지 전문가 시스템의 성능을 향상 시킬 수 있음을 보인 것이다.

참 고 문 헌

[1] J. Seok, S. Cho, "Self-Organizing Feature Map with Binary Reinforcement and Constant Adaptation Gain : For an easier Hardware Implementation", *Proc. ICONIP'94*, vol 2, 966-971, 1994.

[2] I.I.Gihman, A.V.Skorohod, *The theory of stochastic process I*, Springer-Verlag, 1974.

[3] H. Ritter, K. Schulten, "Convergence Properties of Kohonen's Topology Conserving Maps : Fluctuations, Stability, and Dimension Selection", *Biological Cybernetics*, vol 60, pp. 59-71, 1988.

[4] I. I. Gihman, A. V. Skorohod, *The Theory of Stochastic Process II*, Springer-Verlag, 1974.

[5] Yu. Ermoliev, R. J-B Wets (Eds.), 'Numerical Techniques for Stochastic Optimization', *Springer-Verlag*, 1980.

[6] 석 진욱, 조 성원, 홍 윤광, "시불변 학습계수와 이진 강화함수를 가진 경쟁학습 신경회로망에 의한 고장진단 퍼지 전문가 시스템의 자동구축", *Proc '96 FAN 춘계종합학술대회 논문집*, pp 153-158, 1996.

[7] T. S. Chiang, C. R. Hwang, S. J. Sheu, "Diffusion for Global Optimization in R^n", *SIAM J. Cont. and Optimi-zation*, vol 25, No 3, pp 737-753, May 1987.